

bound
Periodical 1288585

Kansas City
Public Library



This Volume is for
REFERENCE USE ONLY

YHAAH 3.1004
YTO 2.0004
OH

PUBLIC LIBRARY
KANSAS CITY

80

WABU 3.000
YTO 2.200
00

THE BELL SYSTEM TECHNICAL JOURNAL

A JOURNAL DEVOTED TO THE
SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL
COMMUNICATION

EDITORS

R. W. KING

J. O. PERRINE

EDITORIAL BOARD

C. F. CRAIG

O. B. BLACKWELL

H. S. OSBORNE

J. J. PILLIOD

O. E. BUCKLEY

M. J. KELLY

A. B. CLARK

F. J. FEELY

TABLE OF CONTENTS

AND

INDEX

VOLUME XXVII

1948

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

THE UNIVERSITY OF
THE STATE OF NEW YORK
LIBRARY

THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XXVII, 1948

Table of Contents

JANUARY, 1948

An Experimental Multichannel Pulse Code Modulation System of Toll Quality— <i>L. A. Meacham and E. Peterson</i>	1
Electron Beam Deflection Tube for Pulse Code Modulation— <i>R. W. Sears</i>	44
Metallic Delay Lenses— <i>Winston E. Kock</i>	58
A Non-reflecting Branching Filter for Microwaves— <i>W. D. Lewis and L. C. Tillotson</i>	83
A Note on a Parallel-Tuned Transformer Design— <i>V. C. Rideout</i>	96
Statistical Properties of a Sine Wave Plus Random Noise— <i>S. O. Rice</i>	109
Noise in Resistances and Electron Streams— <i>J. R. Pierce</i>	158

APRIL, 1948

Microwave Repeater Research— <i>H. T. Friis</i>	183
The Measurement of Delay Distortion in Microwave Repeaters— <i>D. H. Ring</i>	247
Frequency Shift Telegraphy—Radio and Wire Applications— <i>J. R. Davey and A. L. Matte</i>	265
Reflections from Circular Bends in Rectangular Wave Guides—Matrix Theory— <i>S. O. Rice</i>	305
The Approximate Solution of Linear Differential Equations— <i>Marion C. Gray and S. A. Schelkunoff</i>	350
Potential Coefficients for Ground Return Circuits— <i>W. Howard Wise</i> ..	365

iii

1288585

FEB 15 1949

JULY, 1948

A Mathematical Theory of Communication— <i>C. E. Shannon</i>	379
An Aspect of the Dialing Behavior of Subscribers and Its Effect on the Trunk Plant— <i>Charles Clos</i>	424
Spectra of Quantized Signals— <i>W. R. Bennett</i>	446
Analysis and Performance of Waveguide-Hybrid Rings for Micro- waves— <i>H. T. Budenbom</i>	473
Methods of Electromagnetic Field Analysis— <i>S. A. Schelkunoff</i>	487
The Evolution of the Quartz Crystal Clock— <i>Warren A. Marrison</i>	510

OCTOBER 1948

Equivalent Circuits of Linear Active Four-Terminal Networks— <i>L. C. Peterson</i>	593
A Mathematical Theory of Communication (Concluded)— <i>C. E. Shannon</i>	623
Transients in Mechanical Systems— <i>J. T. Muller</i>	657
Maximally-Flat Filters in Wave Guide— <i>W. W. Mumford</i>	684
Transient Response of an FM Receiver— <i>M. K. Zinn</i>	714
Transverse Fields in Traveling-Wave Tubes— <i>J. R. Pierce</i>	732

Index to Volume XXVII

B

- Beam Transmission: Metallic Delay Lenses, *Winston E. Kock*, page 58.
Bennett, W. R., Spectra of Quantized Signals, page 446.
Budenbom, H. T., Analysis and Performance of Wave Guide Hybrid Rings for Microwaves, page 473.

C

- Clock, Quartz Crystal, The Evolution of the, *Warren A. Marrison*, page 510.
Clos, Charles, An Aspect of the Dialing Behavior of Subscribers and Its Effect on the Trunk Plant, page 424.
Communication, A Mathematical Theory of, *C. E. Shannon*, pages 379-423; 623-656.
Crystal Clock, Quartz, The Evolution of the, *Warren A. Marrison*, page 510.

D

- Davey, J. R. and A. L. Matte*, Frequency Shift Telegraphy—Radio and Wire Applications, page 265.
Delay Distortion in Microwave Repeaters, The Measurement of, *D. H. Ring*, page 247.
Dialing Behavior of Subscribers and Its Effect on the Trunk Plant, An Aspect of the, *Charles Clos*, page 424.
Differential Equations, Linear, The Approximate Solution of, *Marion C. Gray and S. A. Schelkunoff*, page 350.
Distortion, Delay, in Microwave Repeaters, The Measurement of, *D. H. Ring*, page 247.

E

- Electromagnetic Field Analysis, Methods of, *S. A. Schelkunoff*, page 487.
Electron Beam Deflection Tube for Pulse Code Modulation, *R. W. Sears*, page 44.
Electron Streams, Noise in Resistances and, *J. R. Pierce*, page 158.
Equations, Linear Differential, The Approximate Solution of, *Marion C. Gray and S. A. Schelkunoff*, page 350.
Equivalent Circuits of Linear Active Four-Terminal Networks, *L. C. Peterson*, page 593.

F

- Field Analysis, Electromagnetic, Methods of, *S. A. Schelkunoff*, page 487.
Filter, A Non-reflecting Branching, for Microwaves, *W. D. Lewis and L. C. Tillotson*, page 83.
Filters, Maximally-Flat, in Wave Guide, *W. W. Mumford*, page 684.
Frequency Shift Telegraphy—Radio and Wire Applications, *J. R. Davey and A. L. Matte*, page 265.
FM Receiver, Transient Response of an, *M. K. Zinn*, page 714.
Friis, H. T., Microwave Repeater Research, page 183.

G

- Gray, Marion C. and S. A. Schelkunoff*, The Approximate Solution of Linear Differential Equations, page 350.
Ground Return Circuits, Potential Coefficients for, *W. Howard Wise*, page 365.

H

- Hybrid Rings for Microwaves, Wave Guide, Analysis and Performance of, *H. T. Budenbom*, page 473.

K

- Kock, Winston E.*, Metallic Delay Lenses, page 58.

L

- Lenses, Metallic Delay, *Winston E. Kock*, page 58.
Lewis, W. D., and L. C. Tillotson, A Non-reflecting Branching Filter for Microwaves, page 83.

M

- Marrison, Warren A.*, The Evolution of the Quartz Crystal Clock, page 510.
Matte, A. L. and J. R. Davey, Frequency Shift Telegraphy—Radio and Wire Applications, page 265.
Meacham, L. A. and E. Peterson, An Experimental Multichannel Pulse Code Modulation System of Toll Quality, page 1.
 Mechanical Systems, Transients in, *J. T. Muller*, page 657.
 Metallic Delay Lenses, *Winston E. Kock*, page 58.
 Microwave Repeater Research, *H. T. Friis*, page 183.
 Microwave Repeaters, The Measurement of Delay Distortion in, *D. H. Ring*, page 247.
 Microwave Transmission: Metallic Delay Lenses, *Winston E. Kock*, page 58.
 Microwaves: Reflections from Circular Bends in Rectangular Wave Guides—Matrix Theory, *S. O. Rice*, page 305.
 Microwaves, A Non-reflecting Branching Filter for, *W. D. Lewis and L. C. Tillotson*, page 83.
 Microwaves, Analysis and Performance of Wave Guide Hybrid Rings for, *H. T. Budenbom*, page 473.
 Modulation, Pulse Code, Electron Beam Deflection Tube for, *R. W. Sears*, page 44.
 Modulation System of Toll Quality, Pulse Code, An Experimental Multichannel, *L. A. Meacham and E. Peterson*, page 1.
Muller, J. T., Transients in Mechanical Systems, page 657.
Mumford, W. W., Maximally-Flat Filters in Wave Guide, page 684.

N

- Networks, Linear Active Four-Terminal, Equivalent Circuits of, *L. C. Peterson*, page 593
 Noise, Random, Statistical Properties of a Sine Wave Plus, *S. O. Rice*, page 109.
 Noise in Resistances and Electron Streams, *J. R. Pierce*, page 158.

P

- Peterson, E. and L. A. Meacham*, An Experimental Multichannel Pulse Code Modulation System of Toll Quality, page 1.
Peterson, L. C., Equivalent Circuits of Linear Active Four-Terminal Networks, page 593.
Pierce, J. R., Noise in Resistances and Electron Streams, page 158. Transverse Fields in Traveling-Wave Tubes, page 732.
 Pulse Code Modulation: Spectra of Quantized Signals, *W. R. Bennett*, page 446.
 Pulse Code Modulation, Electron Beam Deflection Tube for, *R. W. Sears*, page 44.
 Pulse Code Modulation System of Toll Quality, An Experimental Multichannel, *L. A. Meacham and E. Peterson*, page 1.

Q

- Quartz Crystal Clock, The Evolution of the, *Warren A. Marrison*, page 510.

R

- Radio: Transient Response of an FM Receiver, *M. K. Zinn*, page 714.
 Radio and Wire Applications of Frequency Shift Telegraphy, *J. R. Davey and A. L. Matte*, page 265.
 Random Noise, Statistical Properties of a Sine Wave Plus, *S. O. Rice*, page 109.
 Receiver, an FM, Transient Response of, *M. K. Zinn*, page 714.
 Repeater Research, Microwave, *H. T. Friis*, page 183.
 Repeaters, Microwave, The Measurement of Delay Distortion in, *D. H. Ring*, page 247.
Rice, S. O., Reflections from Circular Bends in Rectangular Wave Guides—Matrix Theory, page 305. Statistical Properties of a Sine Wave Plus Random Noise, page 109.
Rideout, V. C., A Note on a Parallel-Tuned Transformer Design, page 96.
Ring, D. H., The Measurement of Delay Distortion in Microwave Repeaters, page 247.

S

- Schelkunoff, S. A.*, Methods of Electromagnetic Field Analysis, page 487.
Schelkunoff, S. A. and Marion C. Gray, The Approximate Solution of Linear Differential Equations, page 350.
Sears, R. W., Electron Beam Deflection Tube for Pulse Code Modulation, page 44.
Shannon, C. E., A Mathematical Theory of Communication, pages 379-423; 623-656.
 Signals, Quantized, Spectra of, *W. R. Bennett*, page 446.
 Sine Wave Plus Random Noise, Statistical Properties of a, *S. O. Rice*, page 109.
 Spectra of Quantized Signals, *W. R. Bennett*, page 446.
 Statistical Properties of a Sine Wave Plus Random Noise, *S. O. Rice*, page 109.
 Subscribers, An Aspect of the Dialing Behavior of, and Its Effect on the Trunk Plant, *Charles Clos*, page 424.

T

- Telegraphy, Frequency Shift—Radio and Wire Applications, *J. R. Davey and A. L. Matte*, page 265.
 Theory, A Mathematical, of Communication, *C. E. Shannon*, pages 379-423; 623-656.
Tillotson, L. C. and W. D. Lewis, A Non-reflecting Branching Filter for Microwaves, page 83.
 Toll Transmission: An Experimental Multichannel Pulse Code Modulation System of Toll Quality, *L. A. Meacham and E. Peterson*, page 1.
 Transformer Design, Parallel-Tuned, A Note on a, *V. C. Rideout*, page 96.
 Transient Response of an FM Receiver, *M. K. Zinn*, page 714.
 Transients in Mechanical Systems, *J. T. Muller*, page 657.
 Transmission: A Mathematical Theory of Communication, *C. E. Shannon*, pages 379-423; 623-656.
 Transverse Fields in Traveling-Wave Tubes, *J. R. Pierce*, page 732.
 Traveling-Wave Tubes, Transverse Fields in, *J. R. Pierce*, page 732.
 Trunk Capacity: An Aspect of the Dialing Behavior of Subscribers and Its Effect on the Trunk Plant, *Charles Clos*, page 424.
 Trunk Plant, An Aspect of the Dialing Behavior of Subscribers and Its Effect on the, *Charles Clos*, page 424.

V

- Vibration: Transients in Mechanical Systems, *J. T. Muller*, page 657.

W

- Wave Guide, Maximally-Flat Filters in, *W. W. Mumford*, page 684.
 Wave Guide Hybrid Rings for Microwaves, Analysis and Performance of, *H. T. Budenbom*, page 473.
 Wave Guides, Rectangular, Reflections from Circular Bends in—Matrix Theory, *S. O. Rice*, page 305.
Wise, W. Howard, Potential Coefficients for Ground Return Circuits, page 365.

Z

- Zinn, M. K.*, Transient Response of an FM Receiver, page 714.

V. 27
Jan - Oct 1948

THE BELL SYSTEM
TECHNICAL JOURNAL

PUBLIC LIBRARY
CITY

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS
OF ELECTRICAL COMMUNICATION

An Experimental Multichannel Pulse Code Modulation System of Toll Quality	<i>L. A. Meacham and E. Peterson</i>	1
Electron Beam Deflection Tube for Pulse Code Modulation	<i>R. W. Sears</i>	44
Metallic Delay Lenses.....	<i>Winston E. Kock</i>	58
A Non-reflecting Branching Filter for Microwaves	<i>W. D. Lewis and L. C. Tillotson</i>	83
A Note on a Parallel-Tuned Transformer Design	<i>V. C. Rideout</i>	96
Statistical Properties of a Sine Wave Plus Random Noise	<i>S. O. Rice</i>	109
Noise in Resistances and Electron Streams...	<i>J. R. Pierce</i>	158
Abstracts of Technical Articles by Bell System Authors...		175
Contributors to this Issue.....		181

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

THE BELL SYSTEM TECHNICAL JOURNAL

*Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.*



EDITORS

R. W. King

J. O. Perrine

EDITORIAL BOARD

W. H. Harrison

O. E. Buckley

O. B. Blackwell

M. J. Kelly

H. S. Osborne

A. B. Clark

J. J. Pilliod

F. J. Feely



SUBSCRIPTIONS

Subscriptions are accepted at \$1.50 per year. Single copies are 50 cents each.
The foreign postage is 35 cents per year or 9 cents per copy.



Copyright, 1948
American Telephone and Telegraph Company

The Bell System Technical Journal

Vol. XXVII

January, 1948

No. 1

An Experimental Multichannel Pulse Code Modulation System of Toll Quality

By L. A. MEACHAM and E. PETERSON

Pulse Code Modulation offers attractive possibilities for multiplex telephony via such media as the microwave radio relay. The various problems involved in its use have been explored in terms of a 96-channel system designed to meet the transmission requirements commonly imposed upon commercial toll circuits. Twenty-four of the 96 channels have been fully equipped in an experimental model of the system. Coding and decoding devices are described, along with other circuit details. The coder is based upon a new electron beam tube, and is characterized by speed and simplicity as well as accuracy of coding. These qualities are matched in the decoder, which employs pulse excitation of a simple reactive network.

I. INTRODUCTION

IN THE development of systems for transmitting telephonic speech, much effort has been directed toward minimizing the effects of noise picked up in the transmission medium. The system described in this paper represents one method which has been successful in eliminating completely such effects under appropriate and practical conditions. In this method, known as Pulse Code Modulation^{1,2,3} (PCM), telephone waves are represented by sequences of on-off constant-amplitude pulses.

Perfect reception of such pulses demands simply recognition of whether any pulse exists or not. Recognition can be carried out effectively in the presence of noise and interference amounting to a substantial fraction of the pulse amplitude. In contrast, telephonic waves carry information by subtle amplitude variations in the course of time. High quality telephone reception accordingly demands a much lower ratio of noise and interference—lower by as much as 50 decibels.

The magnitude of this figure exhibits one good reason for exploring the possibilities of PCM. Another potent reason, which is of particular importance in multi-link transmission, is that, with pulses involving just two standard values, regeneration can be used at repeater points and at the receiver to wipe out transmission impairments. The regeneration process consists of the pro-

¹ A. H. Reeves, U. S. Patent 2,272,070.

² "Telephony by Pulse Code Modulation", W. M. Goodall, *Bell System Technical Journal*, July, 1947.

³ "Pulse Code Modulation", H. S. Black and J. O. Edson; presented June 11, 1947 at the Montreal Summer Meeting of the American Institute of Electrical Engineers, and to be published in the *A. I. E. E. Transactions*.

duction of a properly formed standard pulse, free of noise, to correspond with each received pulse, even though the latter may be considerably misformed. The sole proviso here is that before regeneration the level of noise and distortion in each link be kept below the comparatively large threshold value at which a mark cannot be distinguished from a space. If this holds good throughout the transmission path then literally the received pulses can be made as good as new. In contrast it is impossible fully to repair or to regenerate signals not involving standard values of amplitude and of time. With such signals distortion and noise in each span contribute to the total which therefore increases with the system length.

To sum up, conversion of speech to a code of pulses and spaces permits telephony to assume certain new and desirable properties; ability to work with small signal-to-noise ratios, and ability to regenerate any number of times with no degradation of quality. These advantages do not accrue without certain penalties. Conversion of speech waves to pulse form and back to speech involves a certain degree of apparatus complexity at the terminals. This complexity is not decreased by the need to handle pulses at high speeds, of the order of a million per second. Here radar and television circuit techniques are helpful. Another characteristic is that a greater band width is occupied in the transmission medium. This arises through the operation of two factors, of which one is the use of double sideband in pulse transmission (as against single sideband in carrier telephony), and the second involves the number of pulses used in the code. The relatively wide band required can best be accommodated in the microwave region and it happens that the properties of on-off pulse transmission can be used there to particular advantage.

The PCM system to be described was set up to evaluate experimentally the problems involved in providing multichannel facilities of toll system quality. It was designed to accommodate 96 channels. For experimental studies of such things as crosstalk and methods of multiplexing channels, a fraction of the total number of channels is sufficient and only 24 of the 96 were built. These are arranged as two groups of 12 channels each. The channels of a group are assembled on a time division basis. Assembly of the groups is carried out on a frequency division basis, each group amplitude-modulating its own carrier. In a planned alternative arrangement the group pulses may be narrowed and interlaced to put all 96 channels in time division on a single carrier, but this alternative will not be explored here.

The assignment of 12 channels per group fits in well with the present arrangement of carrier telephone circuits used in the Bell System plant, such as Types J, K, and L.⁴ Use of time division for a group of this size involves pulses with

⁴ "A Twelve-Channel Carrier Telephone System for Open-Wire Lines," B. W. Kendall and H. A. Affel, *Bell System Technical Journal*, January, 1939. "Coaxial Systems in the United States," M. E. Strieby, *Signals*, January-February, 1947.

repetition rates up to 672 kilocycles, and pulse durations as short as a quarter microsecond. These pulses, obtainable from more or less standard types of vacuum tube circuits, can be distributed from point to point in the equipment without too much difficulty. Amplitude modulators and demodulators at two neighboring carriers—65 and 66.5 megacycles—then serve to bring the PCM signals into the intermediate frequency range for transmission to and from the microwave equipment.

The speech quality of the overall system in respect to such factors as band width, volume range, noise, distortion, and crosstalk more than meets the requirements generally imposed upon such systems.

Figure 1 is a front view photograph of the experimental apparatus setup with covers removed from one bay to show typical construction. The two end bays contain intermediate frequency modulators and demodulators required for the two groups. In addition voice frequency terminating sets and jacks are mounted here, together with testing equipment. The center bay and the one to the right of it are identical; each includes apparatus for handling a group of twelve message channels. Transmitting equipment is mounted in the upper half, and receiving equipment in the lower half of each bay. The remaining bay, second from the left, holds all the timing equipment needed to furnish control pulses for operating eight of the message bays, a total of 96 channels. Included are circuits for synchronizing the receiver. Individual regulated power supplies are mounted near their loads on the several bays.

Figure 2 is a rear view of the same equipment. Cables in the four horizontal ducts shown carry control pulses from the timing bay to the 12-channel group bays. These ducts are large enough in cross-section to handle all the cables required for a complete 96-channel terminal.

II. FUNCTIONAL PROBLEMS INVOLVED

The broad problems brought together in building this system may be considered under the four classes following:

1. The pulse code modulation problem; to convert signal waves to pulse patterns.
2. The multiplex problem; to aggregate channels into groups and groups into a supergroup.
3. The transmission problem; to fit the system into the minimum required band width, and to remove the effects of transmission impairments.
4. The pulse code demodulation problem; to convert pulse patterns back to the original signal waves.

These are to be discussed from a functional standpoint to provide background for discussion of the specific equipment.

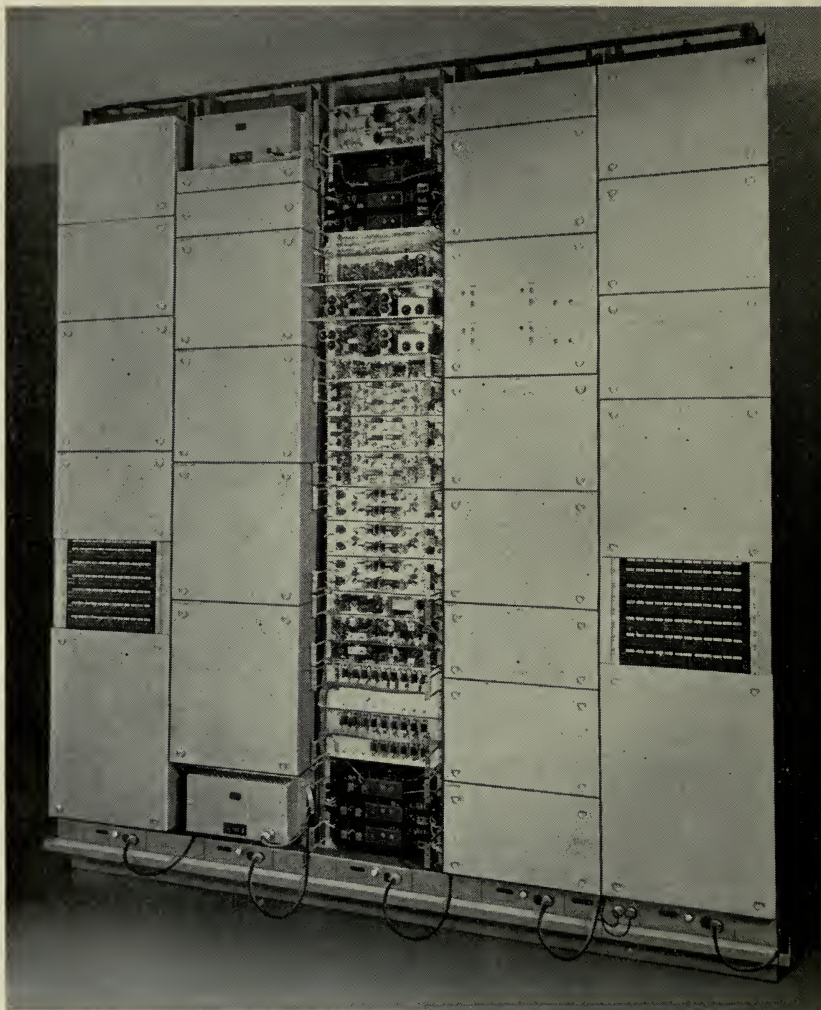


Fig. 1. Front view of experimental PCM terminal equipment, with covers removed from a 12-channel group bay.

Pulse Code Modulation

Sampling. A basic premise of pulse modulation systems is that the information content of a wave can be conveyed by samples taken at sufficiently frequent, equally spaced time intervals. The interval should be no greater than half the period of the highest frequency speech component to be reproduced or, otherwise expressed, the sampling rate should be not less than twice the frequency of the highest speech component present. This provision insures

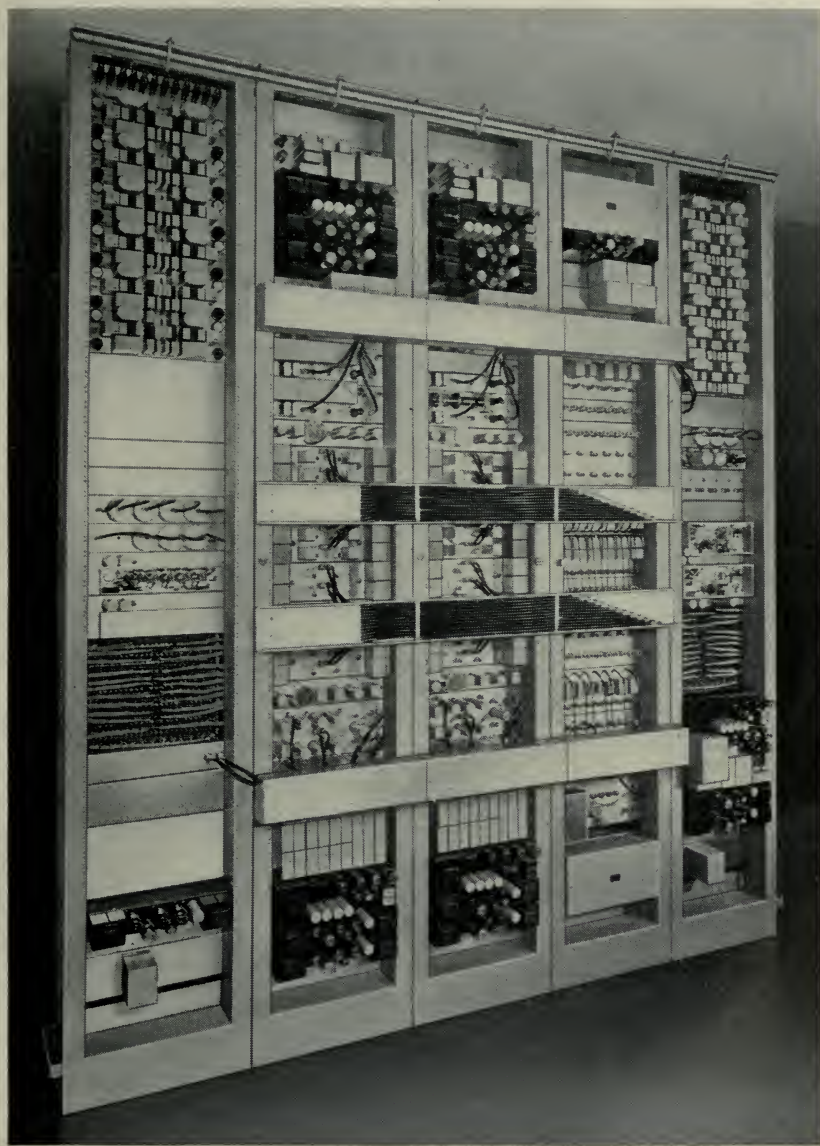


Fig. 2. Rear view. Cables in the horizontal ducts carry timing pulses leftward from the synchronizing bay.

that sidebands produced by sampling do not overlap to introduce distortion, as illustrated in Fig. 3. For a speech band extending up to 3400 cycles, a reasonable sampling rate is eight kilocycles.

Samples may be intermittently-transmitted portions of the signal wave, of appreciable duration, or they may be essentially instantaneous amplitudes.

To afford time for coding, the instantaneous samples may be maintained at constant value for an appropriate interval—a process here referred to as “holding.”

Quantization. The fundamental operation of PCM is the conversion of a signal sample into a code combination of on-off pulses. In any practical system a continuous range of signal values cannot be reproduced since only a finite number of combinations can be made available. Each combination stands for a specific value, of course, so that we wind up by representing a continuous range of amplitudes by a finite number of discrete steps. This process is spoken of as quantization, a quantum being the difference between two adjacent discrete values. Graphically this means that a straight line representing the relation between input and output samples in a linear continuous

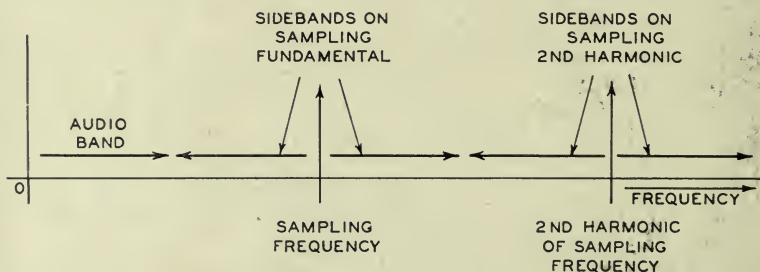


Fig. 3. Spectrum of a sampled audio band, illustrating separation of components when the sampling frequency is at least twice the top audio frequency.

system is here replaced by a flight of steps as in Fig. 4a. The midpoints of the treads fall on the straight line, and the height of the step is the quantum.

Manifestly the greatest error inherent in quantization amounts to half a step. Hence the quality of reproduction may be measured by the size of that interval, which depends upon the total number of steps in the amplitude range covered. With n pulses assigned to represent an amplitude range, the maximum number of discrete steps is 2^n , and the size of each step is proportional to 2^{-n} times the amplitude range.

This error shows up as a noiselike form of distortion, affecting background noise in the absence of speech, and accompanying speech as well. The distortion actually consists of a multiplicity of harmonics and high order modulation products between signal components and the sampling frequency scattered fairly evenly over the audio spectrum. If the audio signal is a simple sine wave, these many products may be identified individually; but for speech or other complex signals they merge into an essentially flat band of noise that sounds much like thermal noise. Since the level of this distortion is fixed by the quantum size, an adequate number of steps must be provided for the lowest

amplitude sounds it is necessary to transmit. Considering that consonant levels may be of the order of 30 decibels below vowels, and that weak talkers may be of the order of 30 decibels down from loud talkers, it is clear that amplitudes as far below maximum as 60 decibels require at least a few steps.

Ordinarily this would involve a large number of pulses for transmission, with a consequent increased complexity of terminal apparatus, and an in-

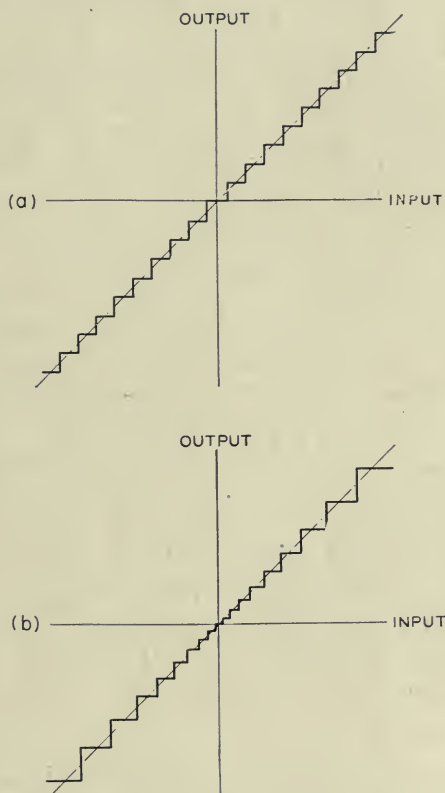


Fig. 4. Relation between input and quantized output, with quantization uniform in (a) and tapered in (b).

creased frequency band per channel. A more practical solution is to employ tapered steps rather than uniform ones. In this way a given number of steps can be assigned in greater proportion to the low amplitudes than to the highs, as shown in Fig. 4b. There results a degree of step subdivision sufficient to care adequately for the low-amplitude sounds including background noise. A small penalty is paid at the upper end of the amplitude scale because of the proportionately smaller number of steps available there, bringing in higher

quantizing noise in this range. How large the effect may be is determined by the degree of taper. To apportion the noise at different levels, the distribution of steps over the amplitude range can be varied.

Preliminary studies of quantization, involving listening tests and noise measurements for various numbers of digits and various kinds of taper, led to the choice of a seven-digit code (128 steps) for the present system. The taper employed reduces the smallest steps 26 decibels below the average size and increases the largest ones about 6 decibels.

Coding. The coder is required to set up a pulse code combination for each quantized signal value. A great many codes are conceivable, but in practice a simple one in which the pulses correspond to digits of the binary number system allows greatest simplicity at the receiver.

While coders may take a wide variety of forms, they can be arranged in three categories according to the way in which they evaluate speech amplitudes. In the first category an amplitude is measured by counting out, with a binary counter for example, the number of units contained in it one by one until the residue amounts to less than a unit. In the second, the amplitude is measured by comparison with one digit value after another, proceeding from the most significant digit to the least, and subtracting the digit value in question each time that value is found to be smaller than the amplitude (or its residue from the previous subtraction). In the third, amplitude is measured in toto by comparison with a set of scaled values. Modulators disclosed by Reeves¹ and by Black and Edson³ are of the first category. That described by Goodall² belongs to the second, and the one described in the present paper is in the third category. Generally speaking the number of operations and the time required for coding decrease in going from the first to the third. Rapid coding is obviously desirable since it allows more channels to be handled in time division by common equipment.

Multiplex

Channels may be multiplexed by arranging them in time sequence, or by arranging them along the frequency scale. These methods are known as time division and as frequency division, respectively. The first is accomplished by gating, or switching, at precisely fixed times. One way of doing this impresses a more or less rectangular pulse on one of the grids of a gate tube, so that a signal wave may be transmitted during the gating pulse. The second method here refers to the use of amplitude modulators, each supplied with an appropriate carrier, which translate the signals to their assigned positions on the frequency scale. To avoid crosstalk in a time-division system, operations in group equipment common to a number of channels must proceed without memory of the amplitudes of preceding channels, requiring build up and decay times to be held within limits. This implies a sufficiently wide pass band

together with phase linearity. For the same purpose in a frequency-division system, filters are used to select and to combine channels. To avoid crosstalk the filters must be sufficiently selective, and amplitude non-linearity must be held within limits.

In the present system, the pulse code is delivered by the coder as on-off pulses in time sequence. It is therefore natural to organize the pulses of the different channels so that they appear in sequence, thus forming a time-division multiplex. Most types of coder require an appreciable length of time, after delivering the pulses of one channel, to prepare for coding the next. This preparation time may be afforded conveniently, without introducing gaps in the pulse train between assignments of consecutive channels, by providing two coders, which take turns at the channels in each time-division group.

As the number of channels increases, evidently the time interval which can be assigned to each channel must be reduced since all of them must be fitted into one period of an 8-kilocycle wave. Similarly the allowable duration of a code or digit pulse becomes shorter as the number of time-division channels in a group is increased. Then too, pulses tend to become more difficult to generate and transmit as their duration decreases. For these reasons it is desirable, and eventually it becomes necessary, to restrict the number of channels included within a time-division group. Frequency division may then be used to aggregate several time-division groups.

For our purposes groups of 12 channels are multiplexed by time division. With seven digits per channel, each group has a capacity of 672,000 pulses per second. To combine eight of the groups for a 96-channel system we again have a choice between frequency division and time division. The equipment of Fig. 1 is laid out to accommodate either procedure. In the first case each group is assigned a carrier for amplitude modulation, as used in actual tests to be described. For the second case the pulse durations would be cut down by a factor of eight, and the pulses from the different groups interlaced. Here the supergroup would have a capacity of 5,376,000 pulses per second.

In carrying out coding operations, and in multiplexing on a time-division basis, various control pulses are required which differ in repetition frequency, in time of occurrence, and in duration. These may be generated from a stable base frequency oscillator through the use of harmonics or, alternatively, of sub-harmonics. In a time-division system which requires a variety of flat-topped waves for switching operations, the use of sub-harmonics fits naturally. Frequency step-down circuits of the multivibrator type produce waves either approximating the desired forms directly, or requiring only simple circuits for reshaping. In contrast harmonic generation requires filters for component wave selection, more and more elaborate in structure as the order of the wanted harmonic goes up. Then, after selection, the harmonic has to be amplified and limited or otherwise shaped for switching purposes. Generally speaking

we need less apparatus and less space if we use multivibrator step-downs. Another advantage is that noise in the base frequency supply produces proportionately less phase jitter with sub-harmonics. While multivibrators do not have high inherent stability, they are capable of great precision when suitably controlled. For these reasons frequency step-downs are used to generate all the timing waves of the system.

Timing and gating operations require accurate time alignment of pulses, which may be accomplished by suitably delaying one set with respect to the other. For this purpose use is made of delay networks or cables, or delay multivibrators. Pulse durations may be varied through the use of interference effects between given pulses and their delayed replicas. Additional timing and gating wave forms are required for regeneration and for assembly in time-division multiplex. All such control pulses are economically generated at a single common point rather than by a number of local generators, and can then be supplied to the message equipment by common power amplifiers via shielded cable.

Transmission

In this section we are to consider the general factors entering into satisfactory reception of on-off pulses including such limitations as those on band width and noise. These are to be viewed while keeping in mind the procedures available for pulse regeneration.

If we start with a rectangular pulse like one of those generated at the transmitter, the corresponding frequency spectrum exhibits lobes extending indefinitely on the frequency scale, with progressively decreasing amplitudes as indicated in Fig. 5a. When such a pulse is passed through a linear phase filter⁵ which discriminates against frequencies beyond the first lobe, the resulting pulse is practically of the sinusoidal form shown in Fig. 5b. Reducing the high-frequency response to the extent prescribed rounds the corners of the transmitted pulse, its duration at half value remaining equal to that of the original input pulse. In practice, transmission characteristics depart from phase linearity and low frequency cutoffs exist. Both effects introduce irregularities into the pulse response. While actual pulses therefore differ slightly in detail from the idealized picture given above, that picture will be retained for simplicity in discussion.

The band width needed for good pulse transmission can be minimized by making pulses as wide as possible. But since a specified number of pulses have to be put into a given time interval (84 pulses in $\frac{1}{8000}$ sec.) we must limit our broadening at a value where the presence or absence of a pulse may be clearly

⁵ A suitable filter is one with a Gaussian characteristic, the loss in decibels varying as the square of frequency and having a value of 1 neper (8.68 decibels) at a frequency equal to $1/T$.

determined in the presence of noise and interference. This spaces the sinusoidal pulses so that they overlap at their half-value points as illustrated in Fig. 5c. There it will be observed that no matter how many pulses occur in succession, the maximum amplitude of the pulse train is no different from that of a single pulse. The spectra of all such pulse trains have a common envelope, shown dashed in Fig. 5c.

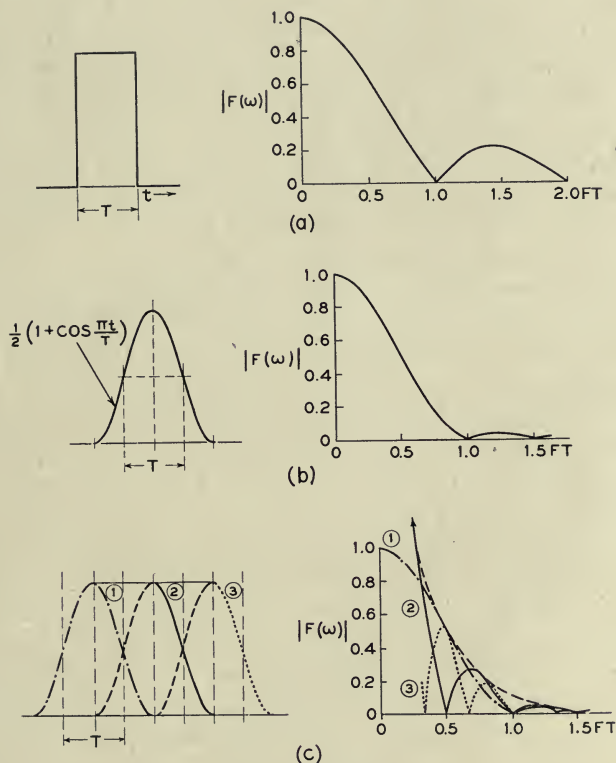


Fig. 5. Pulse forms and their associated amplitude spectra for: a, rectangular pulse; b, single lobe of a sinusoidally varying pulse; c, succession of pulses, each like that in (b).

Further, the band needed to transmit any sequence of such pulses is the same as that needed for a single pulse. This can readily be shown by using the familiar relation between transient build-up time and band width. Moreover, this conclusion is consistent with statistical analysis of all possible pulse combinations, which yields a spectrum of the form of Fig. 5b.

The relation between pulse duration and band width here described gives close to the optimum ratio of signal to noise and interference for the system considered. Narrowing the band would increase build-up and decay times, leading to reduced pulse amplitude and to increased pulse overlap. Thus,

although the narrower band would admit less extraneous noise, margins over noise and interference would be reduced. Widening the band, on the other hand, would allow pulses to build up and decay faster, but would not increase the pulse height. Thus the same signal would result, but the widened band would pass increased noise, and again the margins would be reduced. The optimum band width represents the most useful compromise in efforts to reduce noise and interference and to increase signal.

The filter characteristic we have been discussing is that of the entire link taking in all selectivity inserted between the practically rectangular pulses originally generated at the transmitter and the pulses delivered to the PCM receiver. Filters at both transmitting and receiving terminals of the link are included, the greater part of the overall selectivity being located at the receiver. With about 1.5 microsecond available per pulse at half amplitude, filters spaced 1.5 megacycle apart accommodate the double sideband and keep the interference between groups within tolerable limits.

To establish the presence of pulses we can set up an amplitude threshold equal to half the normal pulse amplitude, and test to see if that threshold is exceeded at a time near the center of the assigned pulse position. Selecting the threshold at half amplitude provides equal margin against the possibility of noise and interference bringing the full pulse amplitude, or mark, below threshold and bringing the nominal space above threshold. Testing at the pulse position midpoint maximizes this margin.

The amplitude threshold is set up by slicing a thin section horizontally out of the pulse at its half-amplitude level by means of an amplitude discriminator. Evidently, this procedure restores the flat top of the pulse. To complete regeneration by restoring the pulse epoch to a standard value the sliced pulses are gated at the midpoints of their proper intervals with narrow pulses supplied by the timing equipment. By these two pulse regenerating processes—slicing and gating—noise and interference are made impotent to produce errors until they attain a substantial fraction of the pulse amplitude. With the effects of noise and distortion thus eliminated, the only noise inherent in the system is that of quantization. In long systems having many repeater points, regeneration has to be practiced at spans short enough to permit cleaning out noise and distortion before it piles up above threshold. Thus in this system transmission impairments have to be considered only for the span between regeneration points; with their effects limited below threshold they are not carried over from one span to the next.

Where PCM groups are multiplexed by frequency division, amplitude non-linearity of the system must be kept within limits. Otherwise intermodulation products may fall within transmission bands, adding to interference. Overlapping of neighboring filter bands also must be kept within bounds to reduce direct crosstalk between the pulse trains. With pulses arranged exclusively

in time-division multiplex, however, amplitude non-linearity of itself is not a factor. In this case the limitation comes on the departure from a suitable attenuation characteristic and from constant delay with respect to frequency. Distortion from these sources may increase the pulse duration so that excessive overlap of adjacent pulses will leave less margin available for interference and noise.

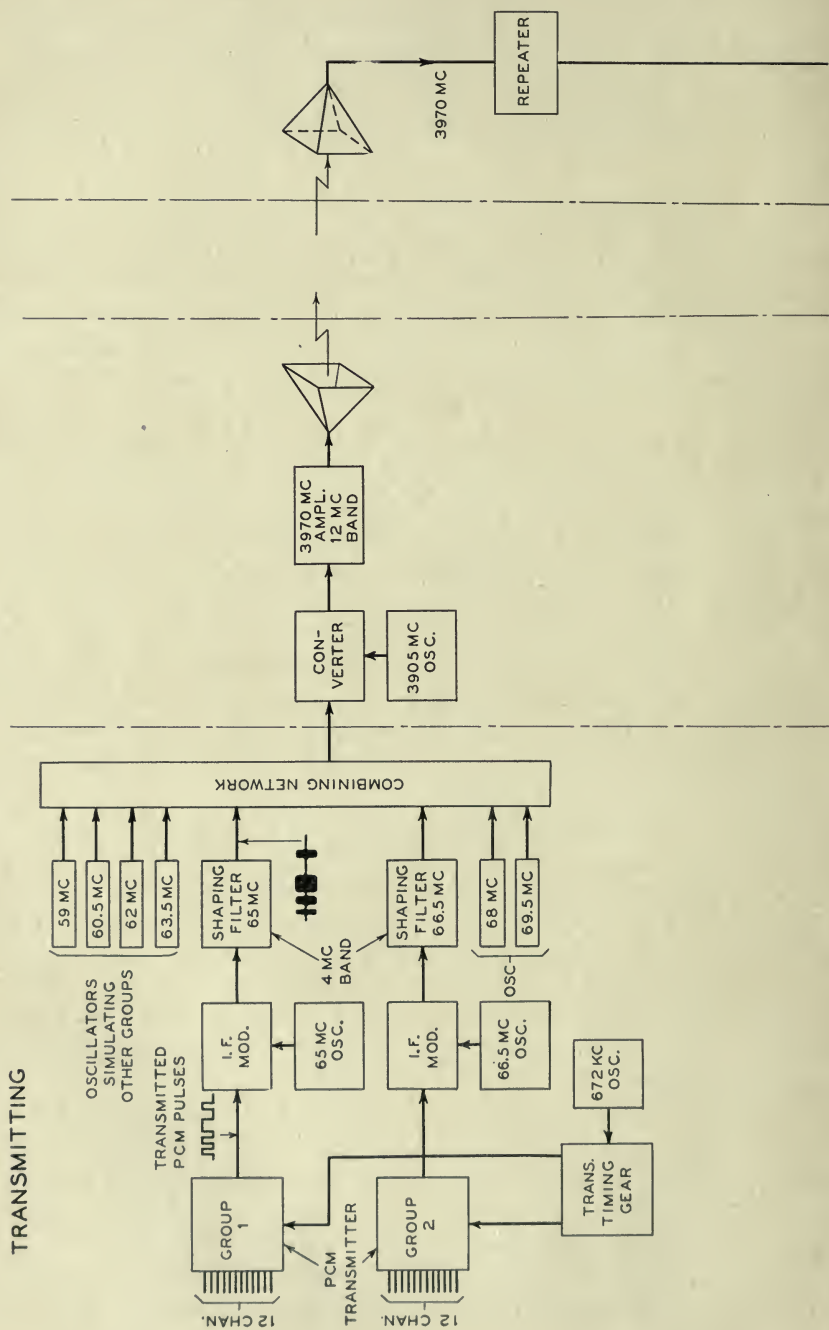
Pulse Code Demodulation

At the receiver, the regenerated code pulses are operated upon to recreate as closely as possible the original signal by operations complementary to those at the transmitter. Unfortunately, however, no process at the receiver can undo the effects of quantization which remain as noise, so that the quanta must obviously be made small enough from the beginning to permit satisfactory speech quality.

In each time-division group alternately working decoders may be used, in the same way as the two coders at the transmitter. Routing is effected by suitably timed gates. Conversion of a pulse code to amplitude may be accomplished by causing each pulse of a code combination to contribute an amplitude corresponding to the binary digit it represents, and then summing the contributions. When tapered steps are employed, due consideration must be given to overall linearity, discussed subsequently. The resulting output is a pulse-amplitude-modulated signal, which is then distributed to the channels of the group by an electronic commutator. Reconstruction of the signal from these distributed pulses is accomplished by simple filtering, which serves to remove components extraneous to speech introduced by the sampling procedure and tied up with the sampling rate.

Production of the necessary timing pulses at the receiving end proceeds in much the same manner as at the transmitter except for one thing. That is, instead of being initiated by a local oscillator, the receiver timing must be linked to the input pulses so that they may be properly routed. This involves the problem of synchronizing or, to borrow a term from television, framing. Use of this term is based upon the similarity of the sequence of PCM digits within a single sampling period to the complete ordered array of television picture elements. Preferably framing should be done with a minimum of time interval and of band width.

One method of synchronizing pulse systems employs a marker pulse which serves to initiate a timing sequence for each frame at the receiver. Here the marker pulse must differ sufficiently from the other pulses to permit its rapid and certain identification. This is ordinarily done by making the marker several times as long as any message pulse. In PCM, however, where digit pulses are run together in many code combinations, the marker would have to be very long to be clearly distinguishable, thereby cutting into channel ca-



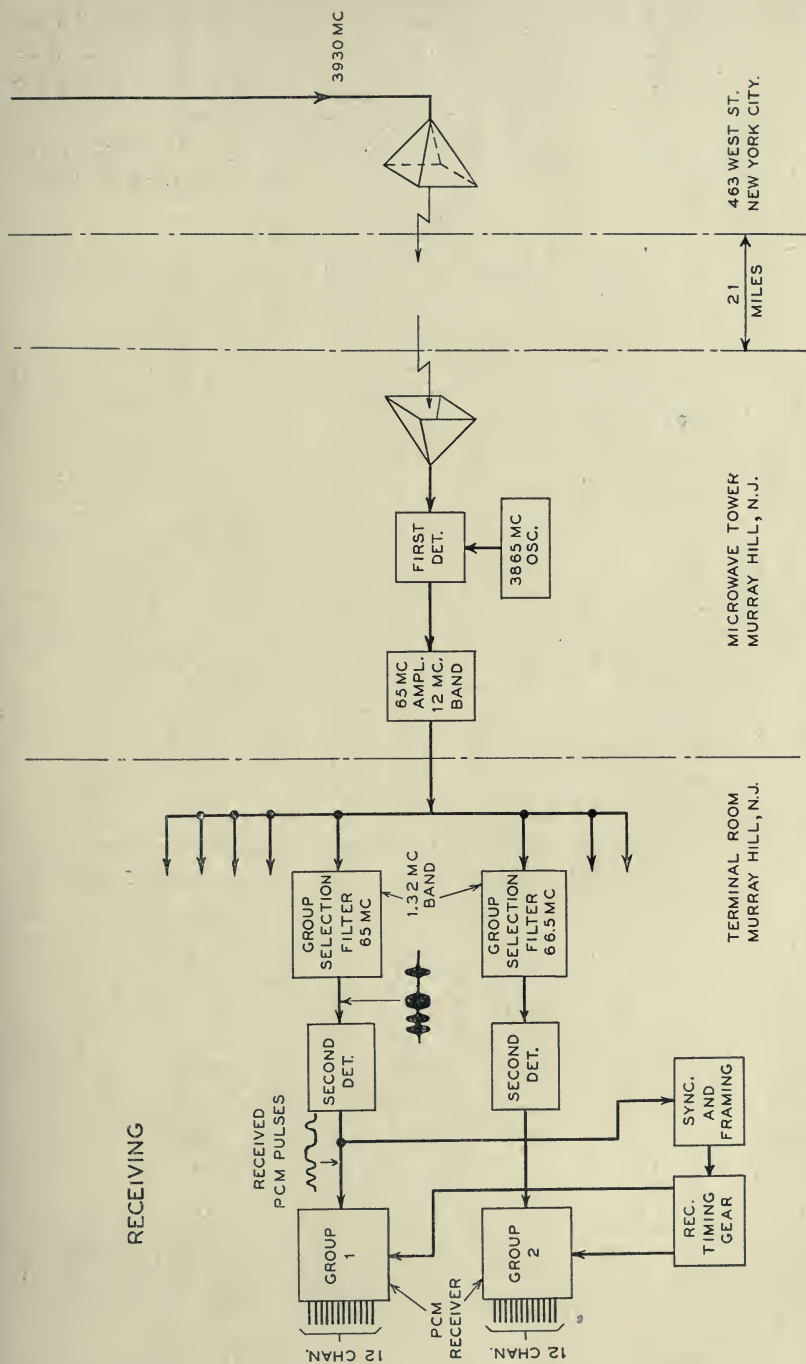


Fig. 6. Block diagram of the experimental PCM system.

capacity. This inefficiency can be avoided by deriving the pulsing rate from the pulse train through the use of a narrow band filter. Framing may then be effected by using in each frame just one digit pulse, which is given a distinctive repetition rate. With this method, a certain amount of time is required to establish synchronism when the system is started up. In the system to be described the framing time is less than one-tenth second—a tolerable value.

III. EXPERIMENTAL SYSTEM

In the light of the foregoing discussion, the block diagram of the experimental system shown in Fig. 6 is believed to be largely self-explanatory. It will be noted that for microwave transmission the modulated intermediate-frequency signals are simply translated in frequency to the 4000-megacycle band. The shaping filters and group selection filters shown have approximately Gaussian characteristics, in accord with transmission considerations noted earlier. The band widths shown for these filters apply between points one neper down from the midband loss. Band widths given for the amplifiers, on the other hand, refer to their regions of essentially flat response.

Overall measurements are facilitated and the amount of experimental apparatus minimized by looping the radio path through a non-regenerative microwave repeater 21 miles away in New York. Both ends of the 24 channels are thus made available at Murray Hill, New Jersey, the location of the apparatus pictured in Fig. 1. With this arrangement, and using conventional 4-wire voice frequency terminating sets, 24 people are able to engage in 12 simultaneous conversations through the system.

PCM Transmitter. The transmitting equipment for an individual 12-channel group is shown schematically in Fig. 7. Each audio input⁶ is passed through a 3400-cycle low-pass filter and through a limiter which chops off the positive and negative peaks of any signal exceeding a prescribed maximum amplitude. This limit is chosen to penalize the loudest talkers to the degree customary in toll system practice. The inputs then enter a "collector" circuit, which assembles samples of the channels in time division multiplex on a common lead. Although it functions electronically under control of pulses from the timing bay, the circuit so resembles a mechanical commutator that this analogy has been used in the schematic. The period of rotation of the "contact arm" is 125 microseconds (8 kilocycles), and a conducting path is formed to the common multiplex lead from each channel circuit in turn for $\frac{1}{12}$ of this period, or $10\frac{5}{12}$ microseconds. It should be

⁶ In telephone terminology, these 4-wire inputs are normally at the -13 decibel level point; i.e., 13 decibels below the transmitting level at the toll test board. Strap connections are provided, as in the Western Electric A-2 channel bank, for adapting the system to inputs 3 decibels smaller. Similarly the final 4-wire outputs are delivered at the +4 (or +7) decibel point. The normal gain through a link is thus 17 decibels but can be set as much as 6 decibels greater.

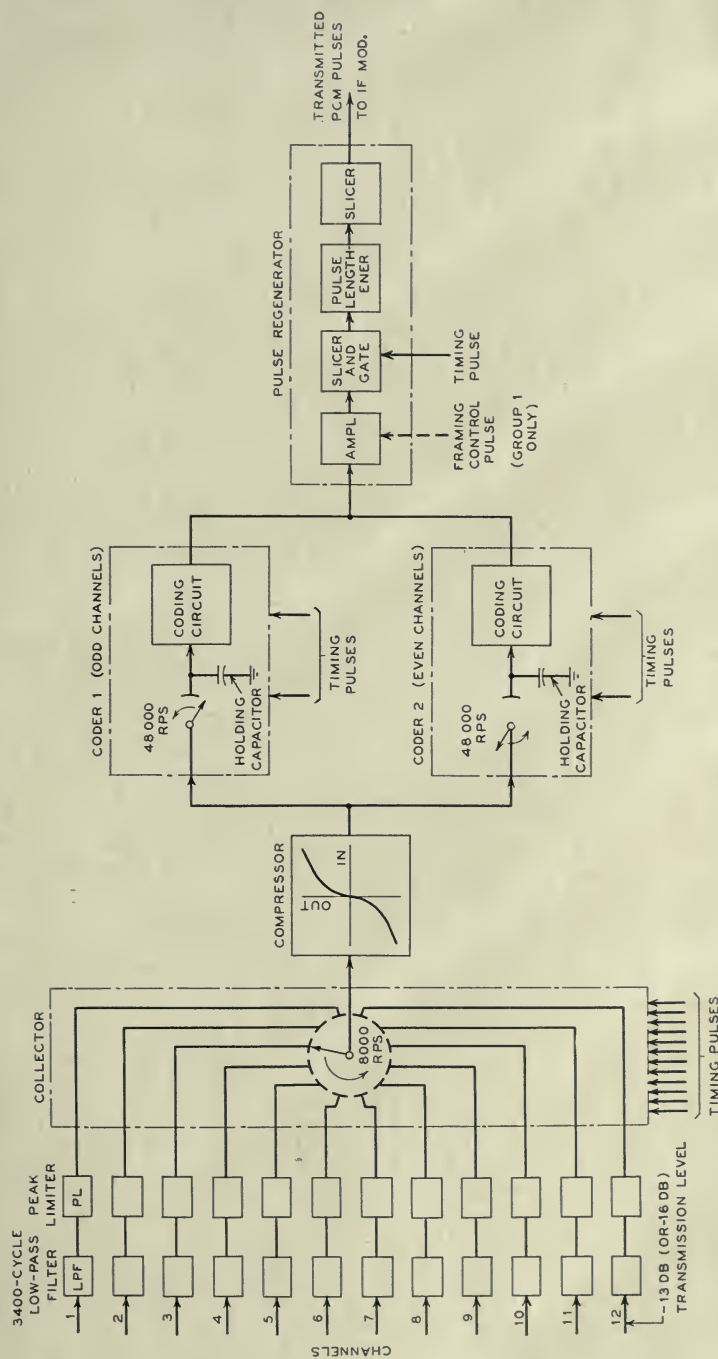


Fig. 7. PCM transmitter for a 12-channel group.

noted that the channel samples thus collected are not "held" at this stage; i.e., each sample does not remain constant in potential during its assigned interval, but rather changes to follow the wave form of the audio signal in the corresponding channel.

The multiplex signal is supplied to an instantaneous compressor, which employs silicon rectifier elements to give an input-output characteristic of the general form indicated within its block (Fig. 7). To understand the purpose of this device, we must recall the discussion of quantizing noise given earlier. There it was found desirable to provide a tapered distribution of step heights in the staircase-like quantizing characteristic, thus devoting a considerable number of small steps to the treatment of background noise and low-level signals. Although coders have been devised which inherently deal with signal amplitudes in this graded manner, it has been found more practicable in the present system to apply amplitude compression to the samples before coding, and to divide this compressed amplitude range into uniform steps in the coder. The result is a tapered step distribution with respect to the original uncompressed scale of amplitudes, details of the distribution being determined by the shape of the compression characteristic.

It may be well to note here that this method can be used in reverse at the receiver, with the decoding performed on an equal-step basis, and the resultant samples passed through a complementary instantaneous expander. If the compression and expansion are truly complementary, the overall characteristic relating amplitudes of input and output samples will be linear except for the tapered array of quantizing steps (Fig. 4b).

Incidentally, no added band width in the transmission path of this system is required to accommodate the instantaneous compandor action.

After compression, the multiplex signal is delivered at low impedance to the inputs of two coders. In Fig. 7 the switch analogy is called upon again to illustrate the routing of alternate samples to the "odd" and "even" coders, and concurrently the storage of these samples on "holding capacitors" to keep them unchanged during the coding operation. Here the switches rotate at 48,000 revolutions per second, each one closing six times in a complete 8-kilocycle frame. The contact segment is drawn as a short arc to indicate a brief closure, actually lasting about 5 microseconds and occurring while the switch of the collector is in contact with a single segment. When the circuit is thus completed from a particular channel to the holding capacitor, the voltage on the latter very rapidly assumes and then follows, for the remainder of the 5-microsecond interval, the potential of the compressed version of the channel signal. When the circuit opens, the latest state of charge, which is essentially an "instantaneous" sample, is left on the capacitor, and is thus held for about 16 microseconds—until the next closure. These sampling operations occur alternately in the two coders.

By a process to be considered later, each coder produces a 7-digit PCM code representation of its set of samples. The two coders deliver their code groups alternately on a common output lead during the final $10\frac{5}{12}$ microseconds of each 16-microsecond holding interval mentioned above. Individual pulses last about 1.5 microsecond, although as delivered from the coders they are somewhat irregular in timing and waveform. It will be noted that the interval allotted to the code group from each channel is just

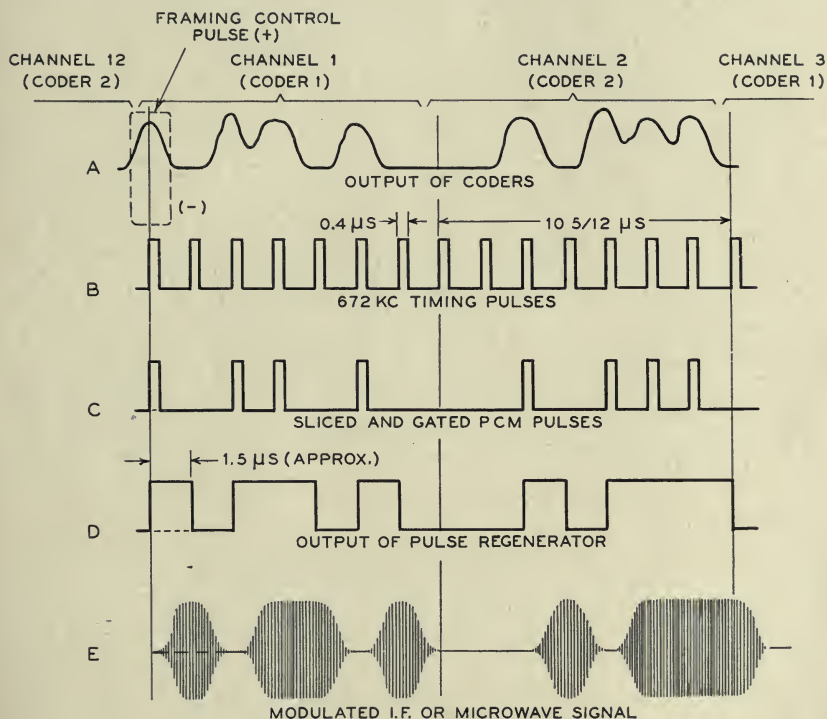


Fig. 8. Waveforms of the pulse regenerator.

$\frac{1}{12}$ of the 125-microsecond frame period and that a continuous train of code groups is thereby produced.

The common output circuit of the coders goes to a pulse regenerator which standardizes the pulses in height by slicing, and in time by gating, as illustrated in Fig. 8. The peaks of the coder output pulses (line A) are lined up in time with the gate control pulses (line B) supplied from the timing gear. The latter have a constant repetition frequency of 672 kilocycles and a uniform pulse length of 0.4 microsecond. Accordingly, the sliced and gated PCM pulses (line C) are also 0.4 microsecond long, and require lengthening to fill their allotted 1.5-microsecond intervals. This is accomplished by a

circuit in the pulse regenerator which first doubles the length of each pulse by adding thereto its own delayed reflection obtained from a short-circuited delay cable, and then slightly less than doubles it again by a similar process using a longer cable. A final slicing, to eliminate amplitude irregularities acquired in the lengthening process, yields square pulses as shown in line *D*, with adjacent pulses merged into a single longer pulse. This is the final output signal delivered to the intermediate-frequency modulator. Passage of the modulator output through a shaping filter results in rounded pulses (line *E*) suitable for transmission over the radio relay path.

In this regenerating apparatus, provision is also made for introducing a "framing control pulse," supplied from the timing bay and normally applied only to Group 1, although any other group may be used if desired. This pulse is about 1.5 microsecond long, and occurs once in each 8-kilocycle frame, but has opposite polarity in successive frames. It is timed to synchronize with the first digit of the Channel 1 code and is large enough in amplitude to override the pulse or space put out by the coder in that position. Hence in the final PCM output from Group 1, pulse 1 of Channel 1 is alternately present and absent regardless of the audio signal. This arrangement, used in automatic "framing" of the receiving timing gear as described in the following section, thus borrows the least significant digit from one channel of the system, leaving that channel usable, but with 6-digit instead of 7-digit quality. A 4-kilocycle tone of very low amplitude which it introduces in that single channel is made inaudible by the low-pass filter in the final audio output.

Synchronization. The connection of a transmitting channel to its proper receiving circuit in the time-division part of the system requires the two terminals to be synchronized: timing operations at the receiver must follow closely those at the transmitter. In a broad and general way this timing matter amounts to getting a local clock to keep the same time as a distant standard clock. Here the criterion of good timekeeping might be thought fussy by some standards; we cannot work with a discrepancy as long as a microsecond for the very good reason that incorrect routing of pulses would then result, associated with intolerably large decoding errors. Three provisions are made to take care of this situation. First, the framing is automatically monitored at all times. Second, if the system is out of frame—as it may be after transmission has been temporarily interrupted—the monitor circuit hunts for and establishes synchronism. Third, whenever the system is not properly synchronized and framed, all message circuits are cut off to avoid resulting noise and crosstalk.

For the purpose of this description we can use a mechanical analogy once more and picture all the transmitting channels of a time-division group arranged in order around a circle (Fig. 9). This time, however, we let each

small division of the circle constitute a commutator bar which is activated by the information of a particular code digit. Thus as the brush is stepped around at a uniform rate it puts digit information on the line in proper sequence. In each complete revolution there appears a set of on-off pulses constituting one frame. With twelve channels in the group, and seven digits to each channel, one revolution of the brush arm covers the frame of 84 digit positions. The revolution period is 125 microseconds. Roughly one and a half microsecond is allotted to each digit, and the brush steps 672,000 times a second. The driving force for stepping is supplied by a

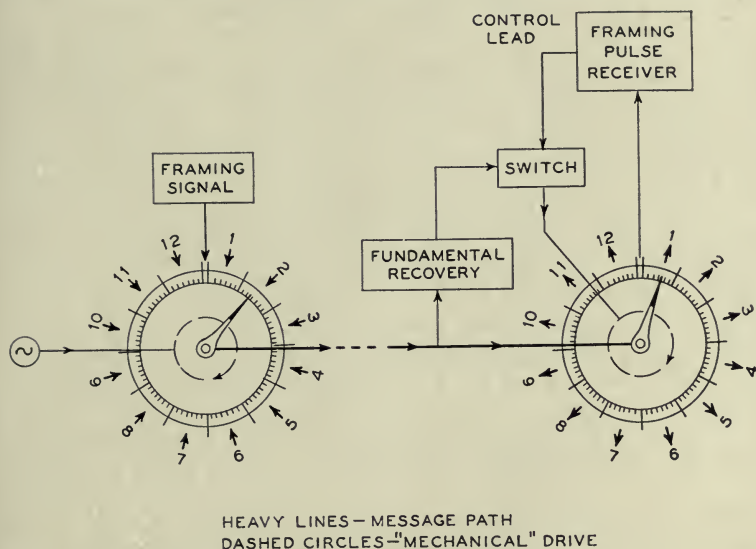


Fig. 9. Representation of the method used for framing, in relation to a single 12-channel time-division group.

stable oscillator, its frequency determined by a quartz crystal. One of the digit positions is taken away from its message duties and assigned for framing purposes, as described in the last section.

Our clock now is a one-handed affair with the brush arm for the hand. Its scale has the customary twelve main divisions, but each main division is divided into seven parts, one for each digit. Hour and minute markings refer to channel and digit, respectively.

At the receiving end we imagine a similar clock structure to be provided. Input pulses go directly to the clock hand which is stepped around to distribute pulses to the twelve message channels, plus a framing channel which takes up but a single digit space. Here the clock has no independent driving source, since it must follow the transmitter. Accordingly, the basic 672-

kilocycle frequency is recovered from the input pulse train by means which include filtering in a narrow-band crystal filter. The filter band is narrow enough to select the base frequency with negligible modulation and noise. That is, sidebands produced by pulse keying are attenuated to negligible proportions, as is the background noise. It is this output which serves to step the brush arm around the frame at precisely the transmitter frequency.

With the transmitter frequency recovered, both hands are stepped around their dials at the same rate. While this is a necessary condition for running the receiver, it is not a sufficient one since most likely the system will not be framed. In fact the odds are 83 to 1 against the two clocks indicating the same time if connection to the receiver is initiated at random times. If we had to deal with ordinary clocks, both in view, the resetting procedure could be accomplished by moving one clock-hand to agree with the other at one fell swoop. But in the PCM case resetting has to be done more discretely since only one dial position per frame is viewed in the framing receiver. Accordingly, an orderly procedure is set up for locating the framing pulse which consists in examining digit positions one by one until the framing pulse is reached. After any one position is viewed long enough to establish the absence of the framing pulse, the receiving clock is set back one digit position and the next position viewed.

This resetting or framing procedure is governed by the framing receiver through its control of a switch which connects the recovered base frequency to the driving mechanism of the clock. If the channels are correctly routed, so that it is the framing pulse which is being viewed by the framing receiver, the switch is left closed, and the 672-kilocycle wave steps the clockhand around the dial without interruption. But if the system is not correctly framed the framing receiver does not get its distinctive pulses. In this case the switch is opened every little while for the duration of a single pulse interval, stopping the local receiver clock during that interval while the transmitter clock advances one digit position. In effect the receiving clock-hand is set back precisely one digit interval with respect to the standard. Thus the next digit pulse is brought into the framing receiver. If again the monitored input turns out to be other than the framing pulse, the stopping process is brought into play once more; this process is repeated until the system is framed.

As pointed out in the last section, the framing pulse is alternately absent and present in successive frames, corresponding to a 4-kilocycle rate. This is readily distinguishable from any of the message pulses, which in practice are found to have little energy content at this frequency. The framing receiver accordingly includes a resonant circuit tuned to four kilocycles. In the hunting process, eight frame periods are allowed between successive interruptions of the clock drive, to give the resonant circuit sufficient time

to build up above a threshold when the system is first framed. With this time interval thus fixed, each clock position is maintained for about one millisecond. Hence the time required to frame the receiver varies between one and eighty-three milliseconds, depending upon the epoch at which the system connection is established.

PCM Receiver. The received PCM signals of a 12-channel group are filtered from the frequency multiplex while in the intermediate-frequency state and then detected to the original signal band, as indicated in Fig. 6. They consist of rounded pulses, nominally sinusoidal in shape, but more or less distorted by transmission defects and accompanied by noise and interference.

These signals are supplied as input to the PCM receiver shown in Fig. 10. They are first sliced in amplitude, the slice being taken at approximately half the average or noise-free pulse height. Code groups of seven pulses are then routed alternately to two decoders, which handle even and odd channels respectively. The routing function is represented in the drawing by a two-segment commutator (*A*) rotating at 48,000 revolutions per second. Before entering the actual decoding circuit, the pulses are again sliced to secure very great uniformity, and are gated with 0.4-microsecond, 672-kilocycle pulses from the receiver timing equipment. Immediately after the arrival of the seventh digit of each code group of these standardized pulses, the decoding circuit produces a voltage on its low-impedance output lead proportional to the quantized amplitude represented by the code. As in the case of coding, details of the decoding process will be reserved for a later section of this paper.

The decoded amplitudes are available only momentarily; therefore it is desirable to sample each one at the proper time and store the result as a charge on a holding capacitor. This sampling process is represented by switch *B* in one decoder and *B'* in the other. Here the switch closures last only two microseconds, and values are held for about 19 microseconds.

The next step is to assemble the six samples from odd channels held successively by one capacitor and the six from even channels held by the other into a single time-division multiplex. Switch *C* performs this operation, rotating at 48,000 revolutions per second, and making contact alternately with the output circuits of the odd and even decoders.

This 12-channel multiplex signal is passed through an instantaneous expander, the purpose of which has been noted in an earlier section. To simplify the problem of making the input-output characteristic of this circuit accurately complementary to that of the compressor, the two devices are designed to use identical silicon units. In the expander, however, the non-linear device is employed in the feedback path of a broadband amplifier rather than in the direct transmission path, thus giving the inverse character-

istic. To allow any compressor to be used with any expander, all the silicon elements are matched to a chosen standard unit, using selection and resistance "padding" methods. By such means, and by use of sufficient loop gain in the feedback amplifier, very satisfactory overall linearity of the system has been attained.

At the output of the expander the waveform of the multiplex signal is essentially the same as that at the input of the compressor in the transmitting terminal. The samples are distributed to their respective channel destinations by a distributor D , resembling the collector described earlier. The rotation rate is 8000 revolutions per second, but the duration of contact on any one segment is only five microseconds instead of the possible full twelfth of the 125-microsecond frame period. This effective narrowing of the contact segments is done to allow the closure to occur well within the interval in which the circuit is completed by switch C from the output of a particular decoder. Each of the 12 segments of the distributor is provided with a holding capacitor, which stores its allotted samples for the full 125-microsecond frame period. The potential on any one of these capacitors thus changes at 125-microsecond intervals from one quantized sample amplitude to the next derived from the same original speech wave.

This potential is of sufficient magnitude to require use of only a simple single-stage triode amplifier for the output of each channel. Lengthening the samples by holding, as described, helps to make this possible by causing the amplifier to deliver useful power continuously instead of on a fractional time basis.

The only disadvantage of using lengthened pulses arises from an effect, very similar to the "aperture effect" encountered in sound movies, which introduces a curving slope across the audio gain-frequency characteristic of the system. In the present case the gain drops about three decibels as the frequency goes from the lowest to the highest value of interest. This slope can be corrected by a simple equalizing network, as shown in Fig. 11. In the present system the equalization is incorporated in the low-pass filter at the input of each audio channel. This is preferable to equalizing at the output, where power is at a premium.

The outputs from the channel amplifiers are passed through 3400-cycle low-pass filters, identical with the input filters except for omission of the equalization, and are delivered to standard voice-frequency circuits at the same levels⁶ as are provided by a type J , K , or L carrier system.

IV. COMPONENT CIRCUITS

Many of the circuit techniques used in the experimental system are conventional, others are more or less unfamiliar, and still others are believed to be novel. In the following some of the more important building blocks are

described. These include sampling circuits, the instantaneous compressor, slicers, and the PCM coder and decoder.

Sampling Circuits. The function of opening and closing a circuit at prescribed instants, represented by rotating switches in the block schematics, is actually performed⁷ by one or the other of the devices shown in Fig. 12, employing diodes and triodes, respectively.

The diode type (Circuit a) is normally biased "open" by rectified charges stored on the two large capacitors. While in this condition it presents an extremely high series resistance between the low-impedance input and the load. But when a flat-topped pulse is impressed upon the pulse transformer, the aforesaid high impedance changes to a low value (of the order of 100

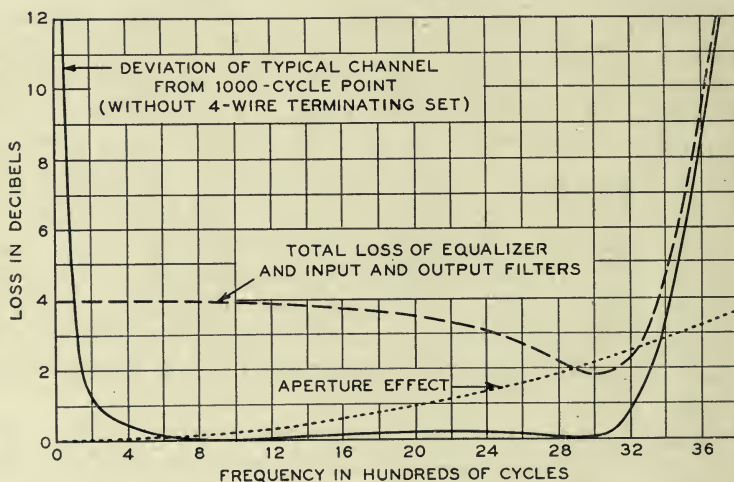


Fig. 11. Equalization for the aperture effect.

ohms), for the duration of the pulse. The upward-sloping tops of the pulses sketched in Fig. 12 represent predistortion to allow for transmission through the pulse transformer, for the purpose of obtaining rectangular pulses at the tube itself.

In the other variety (Circuit b) the plate-cathode paths of the two triodes are connected directly in parallel, conducting in opposite directions. The grids are both arranged to be biased below cut-off during the "open" condition of the sampler by grid rectification of pulses, and to be driven strongly positive by a pulse when a low-resistance conducting path is required between source and load.

These two types have their respective advantages. For example, the low capacitance to ground which the diode type affords across its load makes it

⁷ A single exception is the case of switch A of Fig. 10. In this case two gated slicer circuits are used, as described subsequently.

preferable for use in the collector, where twelve samplers are multiplexed to a common load. On the other hand the triode type affords a d-c. path (without the series blocking capacitors that are required by circuit a) between source and load. In cases where a holding capacitor is to be charged to a succession of sample amplitudes which must be kept mutually independent to avoid crosstalk, the d-c. path is very desirable for it avoids "memory" effects associated with passage of the charging currents through the blocking capacitors. A further useful property of the triode circuit is its ability to

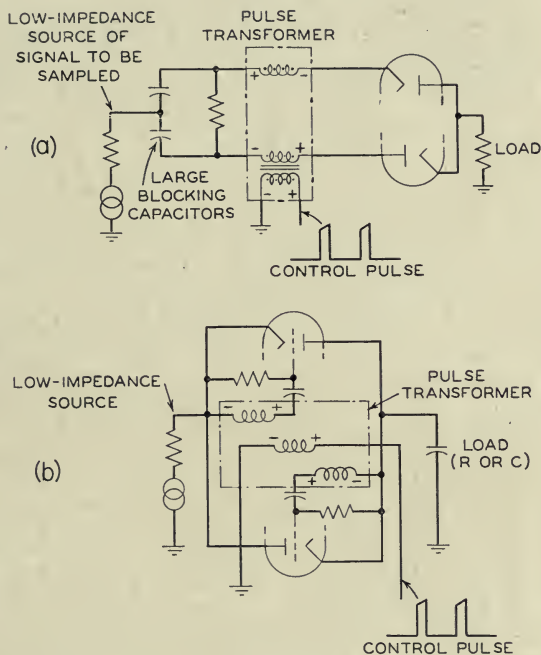


Fig. 12. Diode and triode sampling circuits.

sample signals of greater amplitude than that of the control pulses. With the diode circuit, the signal amplitude must be kept less than half the pulse height.

Instantaneous Compressor. The simple configuration of the non-linear circuit used in the compressor and in the feedback path of the expander is shown in Fig. 13. Two selected silicon rectifiers are connected in parallel, but poled oppositely, with a small padding resistor (R_1 or R_2) in series with each. A parallel padding resistor (R_3) of large value is also provided. The direct-current resistance of this combination varies from about 6000 ohms at zero signal to about 190 ohms at the peak of a full-load signal. Input is applied as a current through the relatively high resistor R_4 , and the voltage

drop across the varistor unit constitutes the compressed output. Although the development of this compressor has been a problem of many interesting aspects, it must suffice here to point out that copper-oxide elements are unsuitable at the speeds involved because of excessive capacitance, that silicon is superior to at least the presently available types of germanium

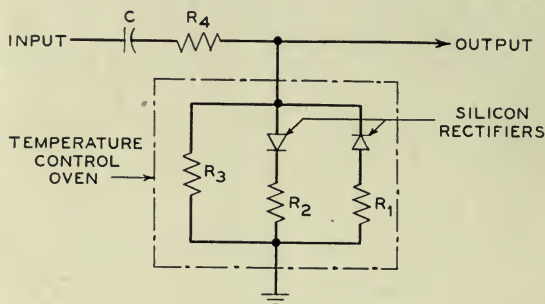


Fig. 13. Instantaneous compressor.

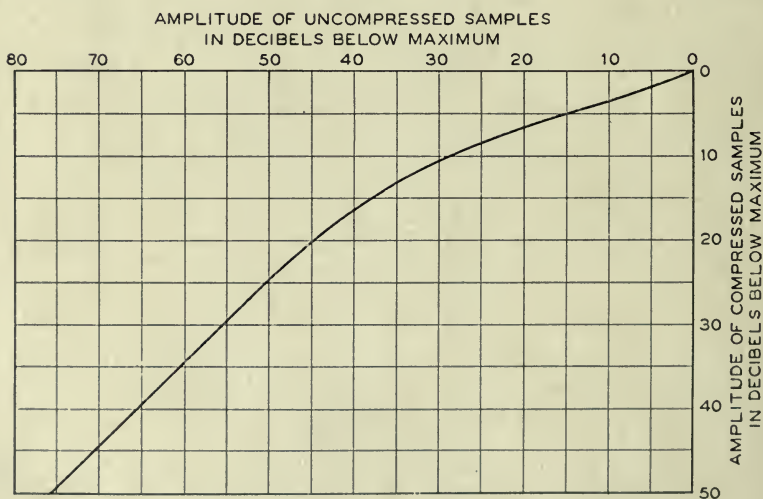


Fig. 14. Characteristic of the instantaneous compressor.

varistor in regard to freedom from certain memory or hysteresis effects believed to be thermal in origin, and that a satisfactory yield of matched silicon elements appears to be obtainable by the selection and resistance padding methods employed. Temperature control is used for the sake of constancy of the compression characteristic. Aging has been found negligible over a period of many months.

The actual compression characteristic afforded by these units is plotted

in Fig. 14 in terms of compressed signal amplitude vs. uncompressed signal amplitude, both in decibels.

Slicers. The slicer circuit shown in Fig. 15 resembles a conventional single-trip multivibrator in configuration, but functions somewhat differently because of the choice of parameters. In particular, capacitor C is made large enough so that the potential drop across it does not vary appreciably during normal operation, and plate resistor R_1 is given a small value such that the gain around the feedback path of the circuit is approximately unity when both triodes are in their active regions. Germanium varistors VR_1 and VR_2 maintain desired bias conditions regardless of the number of pulses or spaces in the input signal. Unlike the single-trip multivibrator, which when tripped remains so until the charge on C has had time to relax,

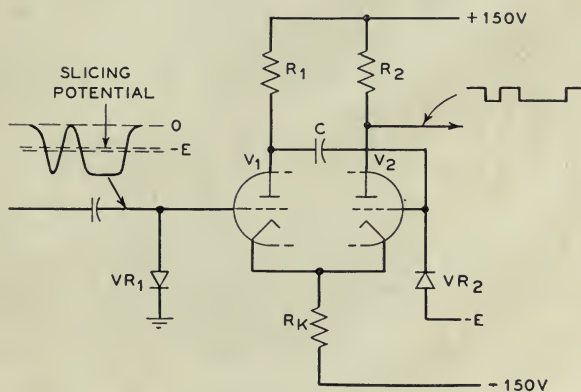


FIG. 15. Slicer circuit.

this circuit trips whenever the input signal falls through a narrow potential range near the value E , and trips back again when the input rises through the same potential range. A square wave of constant amplitude representing an accurate thin slice of the input is thus made available across resistor R_2 . The circuit is capable of high speed, slicing pulses as narrow as 0.1 microsecond.

Time gating may readily be included among the functions of this circuit, through the addition of a triode having its plate and cathode connected directly to the plate and cathode, respectively, of tube V_1 . The grid of this added triode is held normally at ground potential, and is pulsed in the negative direction by approximately rectangular gating pulses, having an amplitude of about $2E$. The total current passing through the common cathode resistor R_K is essentially constant, and if either or both of the paralleled tubes are conducting at a given instant this current is carried by one or shared by both of them. Hence the tripping action, involving transfer of

the cathode current to V_2 , cannot occur as long as the grid of either V_1 or the added tube is positive with respect to the critical slicing potential. Tripping does occur whenever this limitation is removed, and thus the desired gating and slicing functions are performed concurrently.

In the experimental system this circuit is used to regenerate the PCM pulses at the common output of the two coders, and again at the input of each decoder to slice the received pulses and to sort out code groups of odd and even channels.

Coders. The method of binary coding used in this system was originally suggested by F. B. Llewellyn. It employs a novel electron beam tube. This device, pictured in Fig. 16, carries out simultaneously the two functions of quantizing and of coding. The tube is about $10\frac{1}{2}$ inches long and $2\frac{1}{4}$ inches in diameter. It has the 128 combinations of the 7-digit code laid out permanently as holes in an "aperture plate," and translates sample amplitudes from the form of beam deflections directly into PCM pulse symbols. Figure 17 shows one of these tubes in its socket on the rear of a coder panel, and above it a rectangular permalloy shield which covers the coding tube of the other coder serving the same 12-channel group.

As shown schematically in Fig. 18 the coder includes, in addition to the coding tube, a sampling and holding circuit which sorts out the odd (or even) channels from the input multiplex signal, push-pull amplifiers for vertical and horizontal beam deflections, and simple arrangements for blanking, focusing and centering. Within the tube are shown, from left to right, a conventional electron gun, vertical and horizontal deflection plates, a rectangular "collector" for secondary electrons, a "quantizing grid," the "aperture plate" and finally a "pulse plate." Figure 19 shows the target end of the tube, as viewed from a point near the gun. "Digit holes" in the aperture plate, laid out in accordance with the desired binary code, may be seen behind stretched parallel wires of the quantizing grid. One may count 64 narrow holes separated by equally narrow bars of metal in the left-hand vertical column, 32 holes in the next column, 16 in the next, and so on for seven columns. There are 129 grid wires uniformly spaced and accurately aligned so as to mask the upper and lower edges of every one of these holes when viewed from the geometric "point of origin" of the beam.

Stored audio samples from the sampling and holding circuit provide potential for the vertical deflection, with zero at the center, positive amplitudes in the upper half, and negative amplitudes in the lower half of the target area. A sawtooth sweep provides the horizontal deflection. The beam is blanked while deflection potentials are being changed to move it upward or downward from one sample amplitude to the next. When first restored, the beam strikes in the left-hand unperforated region of the aperture plate, and is then swept linearly across from left to right. Electrons which pass through the

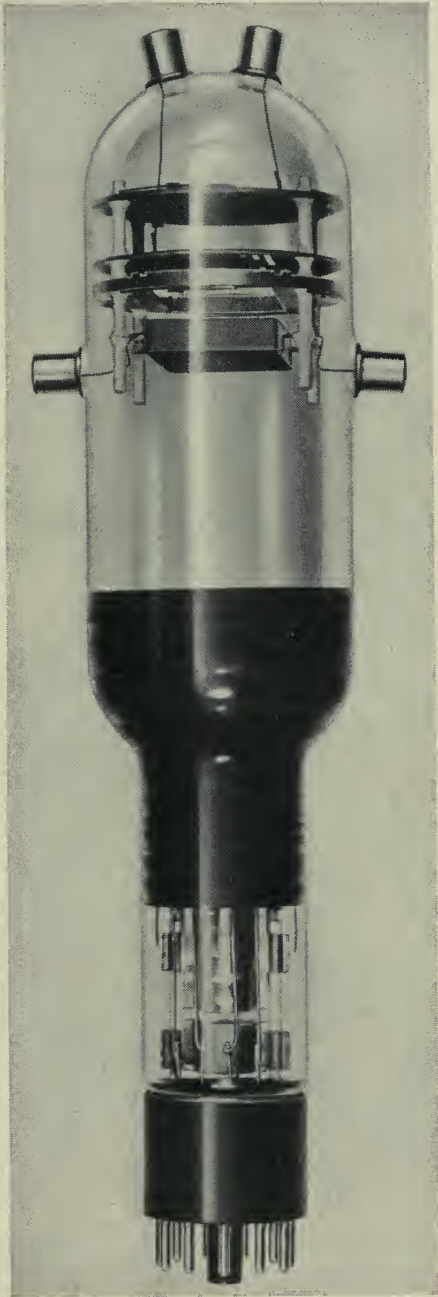


Fig. 16. Coding tube. This new electron device transforms speech samples into pulse codes.

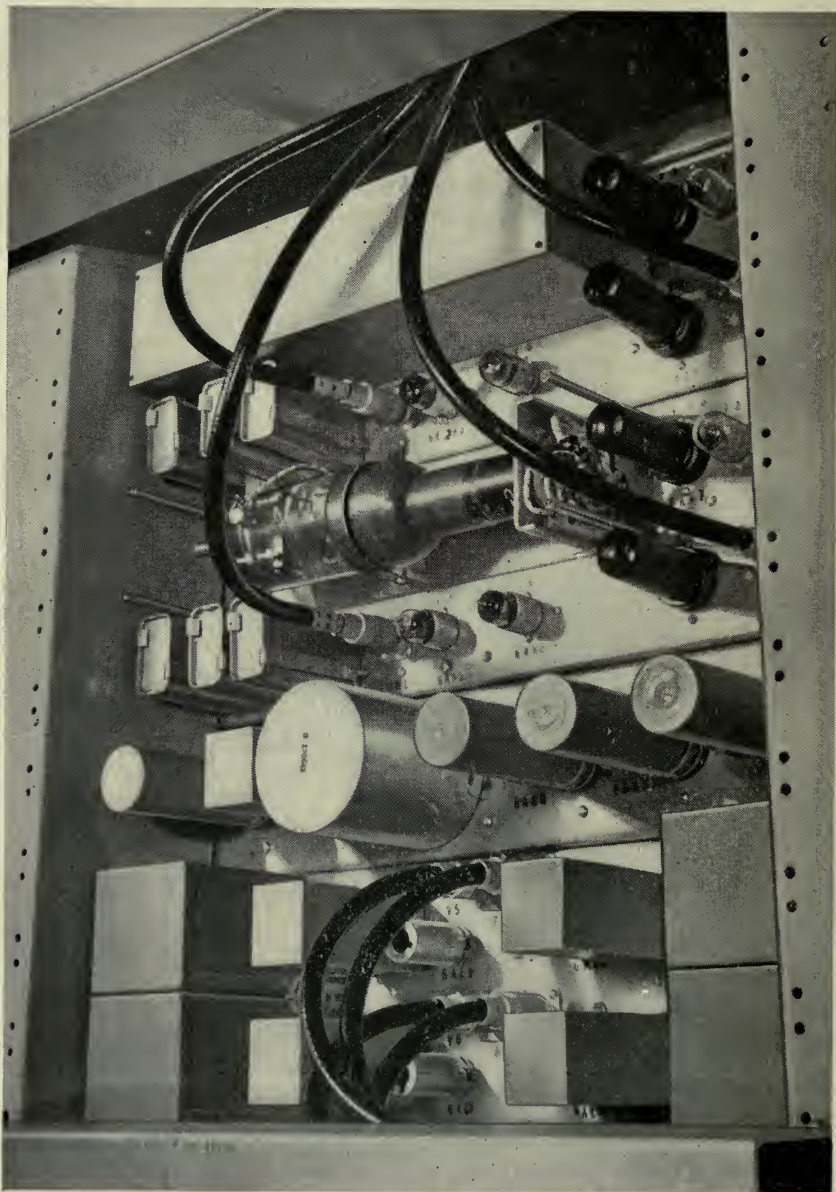


Fig. 17. Rear view of part of the collector (bottom panel), compressor, and two coders.

digit holes during the sweep are caught by the pulse plate, forming pulses which are amplified, gated and lengthened in the pulse regenerator to constitute the desired PCM signals. Retrace of the sweep occurs while the

beam is blanked, and is simultaneous with the application of the succeeding audio sample.

The wires of the quantizing grid are used to guide the beam so that it can illuminate only the particular row of apertures which correspond to the initial vertical deflection. Without this feature, erroneous codes would be produced when the beam straddled the edges of apertures or crossed from one amplitude level into another, as a result of electrode misalignment or possible slight drift in potential of an applied sample. The guiding action (basically proposed by W. A. Marrison and applied to the PCM coder at the suggestion of G. Hecht) is obtained by means of feedback from the quantizing grid to the vertical deflection amplifier.

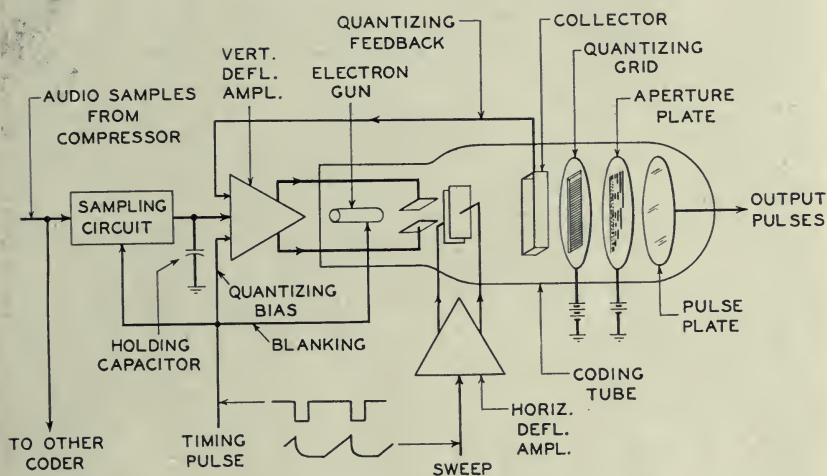


Fig. 18. Functional schematic of the coder.

The feedback signal is actually a current taken from the positively biased collector, which draws to it secondary electrons from the grid wires. The portion of the beam current striking the grid varies as a cyclic function of the vertical deflection. It follows that for some spot positions the value $\mu\beta$ in the feedback loop is positive, for others negative; hence with proper amplifier design there is a stable and an unstable region associated with each wire.

The spot can come to rest (vertically) only within one of the stable regions. In order to locate it consistently near the center of such a region, and thus gain equal margins against "hopping" upward or downward across a wire, a "quantizing bias" is introduced into the vertical deflection amplifier, along with the feedback and the signal samples. This bias is a current of opposite polarity from the unidirectional feedback current, and of magnitude equal

to the average between the two values of feedback current which exist when the beam falls (1) directly on a grid wire and (2) midway between two wires. One may regard this bias as pressing the beam upward against a wire, resisted by downward pressure associated with the feedback current. The latter

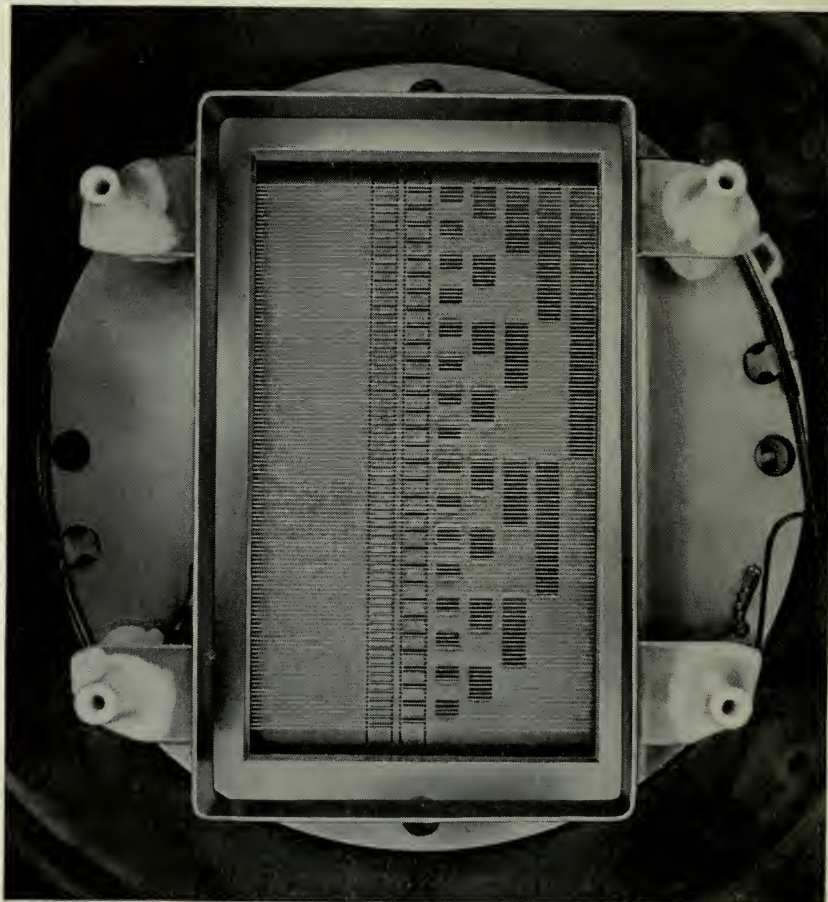


Fig. 19. Interior of the coder tube, viewed from the gun end.

increases as the beam approaches the wire, and equilibrium is reached when the beam is about half way between positions (1) and (2) mentioned above. The feedback current is actually intermittent, turned off and on by the blanking pulse, and careful analysis shows it necessary to make the bias intermittent also, with its wave fronts synchronous with those of the feedback current. This is readily accomplished by deriving the bias from the blanking signal itself.

With this arrangement the beam, suddenly turned on, moves either upward or downward from its initial "unquantized" position to the nearest position of stable equilibrium. Quantization is completed in less than a microsecond. Thereafter, as the beam is swept horizontally across the target area, it remains pressed upward against the lower surface of its guiding wire. Quantization is thus maintained until the end of the sweep, when blanking occurs. The margins against hopping across wires, while quanti-

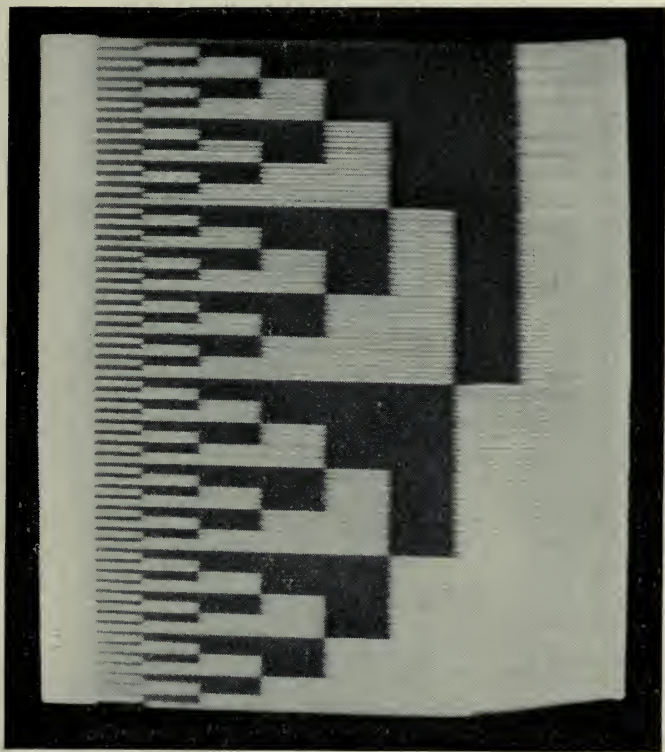


Fig. 20. "Television picture" of the aperture plate of a coder tube.

zation is in effect, are related to the amount of loop gain in the feedback path. In the present system feedback measuring about 20 decibels is provided, and will counteract changes which, without feedback, would move the beam two grid wires in either direction from the initial position.

This coding process, based on the electron beam coding tube and making use of a two-dimensional permanent layout of the code, is more straightforward than the various coding processes which depend upon counting or sequential comparison. Accordingly it leads to higher speeds and greater circuit simplicity.

Figure 20 illustrates the coding accuracy obtained. This photograph of

the screen of a test oscilloscope may be thought of as a television picture of the aperture plate of the coding-tube. To produce the pattern an audio-frequency sawtooth wave of full-load amplitude was applied to the sampling and holding circuit at the input of a coder. The resulting sample amplitudes, quantized by the coding tube and falling into all the possible 128 steps of the quantizing characteristic with uniform regularity, were used to energize the vertical deflection of the test oscilloscope. An ordinary synchro-

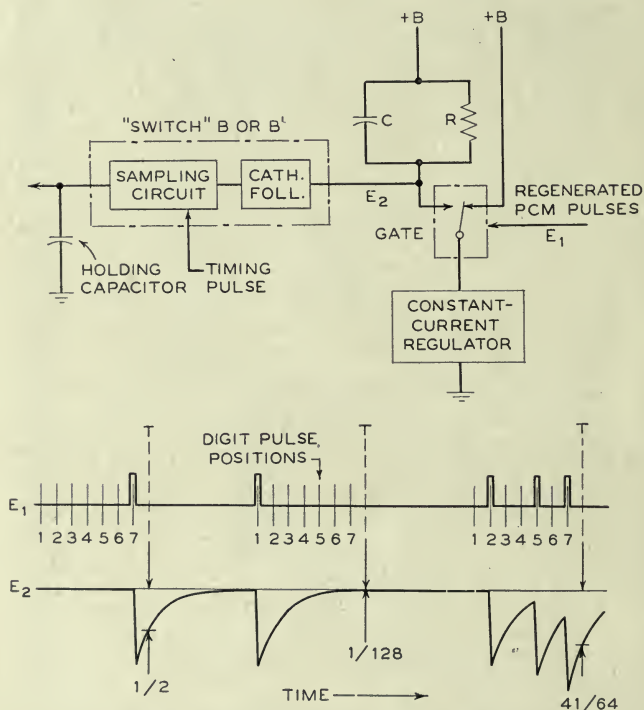


Fig. 21. Shannon decoding circuit and waveforms.

nized sweep provided the horizontal deflection. The square PCM pulses, delivered by the coder via the pulse regenerator, were applied to the intensity control.

Thus the code pulses corresponding to each quantized amplitude were made to appear as a row of blanks in a horizontal trace at the proper relative height. This pattern is very useful in studying coder performance.

Decoders. The decoding method is an impressively simple one originally proposed by C. E. Shannon. In its basic form it employs a pulsed resistance-capacitor circuit as illustrated in Fig. 21. Upon arrival of each pulse of the code, an identical increment of charge is placed upon the capacitor of the

device. The time constant $t = RC$ is such that, during any single pulse interval, whatever charge is on this capacitor decays precisely 50% in amplitude. It follows that the charge remaining at some chosen instant after the arrival of a complete code group consists of contributions of all its pulses, weighted in a binary manner. That is, if we define the contribution of a pulse in the final digit position as $\frac{1}{2}$, then contributions of $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$ and $\frac{1}{128}$, respectively, are made by pulses in successively earlier positions. Any value from 0 to $\frac{127}{128}$, in steps of $\frac{1}{128}$, may thus be produced. Of course the digit holes in the aperture plate of the coder are laid out to make this straight-forward scheme workable. Since samples of low-level audio signals are coded near the center of the aperture plate, the corresponding decoded values lie in the neighborhood of $\frac{1}{2}$.

The basic Shannon decoder, then, comprises the resistance-capacitor circuit, means for supplying it with precisely controlled units of charge at precisely determined times, and a sampling and holding circuit (represented by switch B or B' in Fig. 10) to measure and store the decoded potential which is fleetingly present across the capacitor at a regularly recurring instant T , following the final pulse position. The scheme employed to inject the identical charges involves a regulated source of current and a gate to admit this current to the resistance-capacitor circuit under control of the regenerated PCM pulses. Two successive slicing operations and careful gating, as described earlier, make these pulses more than adequately uniform.

The wave form sketches of Fig. 21 show three typical decoding cycles. In the first, a single pulse in digit position 7 produces a decoded amplitude of $\frac{1}{2}$; in the second, a pulse in position 1 gives $\frac{1}{128}$; and in the third, pulses at 2, 5 and 7 provide a decoded value of $\frac{41}{128}$. It may readily be verified that the provision of an idle channel interval between operations (following from the alternate use of two decoders) allows the residue of one decoding operation to decay to a negligible value (never larger than $\frac{1}{128}$ of a single step height or "quantum") by the time of the next consecutive sampling. Experimentally, interchannel crosstalk from this source is virtually nonexistent.

In the foregoing it has been emphasized that precise timing is required for this basic Shannon decoder. Although the necessary accuracy was actually obtained without great difficulty in early tests, a modification has also been introduced which eases the requirements to a very marked extent. This scheme, devised by A. J. Rack, employs a damped resonant circuit in conjunction with the resistance-capacitance elements in a manner illustrated by Fig. 22. The natural frequency of the resonant circuit L, C_2, R_2 is made equal to the PCM pulse rate, and the time constant of the damped oscillation ($t = 2R_2C_2$) is matched to that of the circuit R_1, C_1 . The same charging

pulses pass through both sections of the circuit; hence by proper choice of C_1 and C_2 the amplitudes of the damped sine wave and the exponential may be proportioned so that the rate of change of their combined potential becomes zero at successive points one pulse period apart. In fact it has been found possible to make both the first and the second time derivatives of potential equal to zero simultaneously at such points.

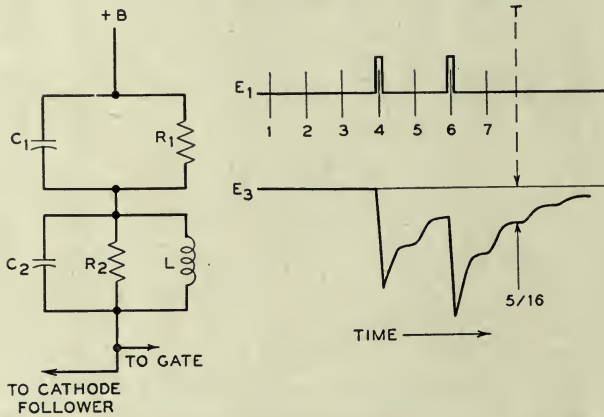


Fig. 22. Shannon-Rack decoder.

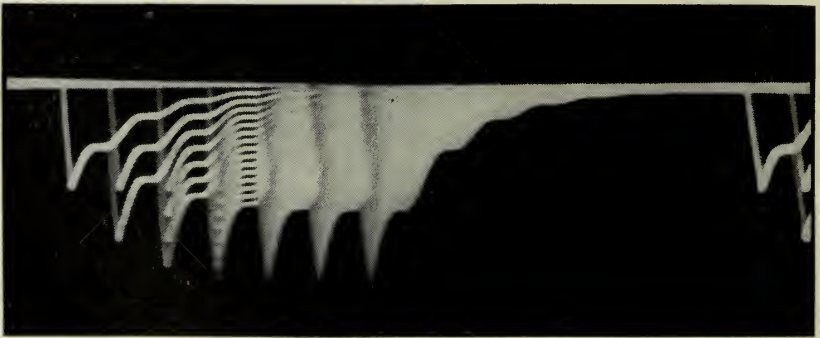


Fig. 23. Output of Shannon-Rack decoder for a signal giving 100% modulation.

By this modification, not only the times of application of the charges but the time of sampling is made much less critical. Of course this presumes that the sampling circuit is designed to complete its operation near the center of a level region. In Fig. 22, the voltage transient due to a typical pair of pulses is sketched, and Fig. 23 shows an oscilloscope screen, on which the waveforms delivered by the cathode follower of one of the Shannon-Rack decoders of the system are superimposed for a full set of 128 possible code combinations in sequence.

V. PERFORMANCE

In general the behavior of the system has shown promise for toll plant application. Among other things the stability realized in the adjustments of the coder and decoder, and the apparent absence of aging or other drift in the compressor and expander have been gratifying. A daily check of the focus and centering in the coders and of the time constants in the decoders appears adequate to keep them in optimum adjustment. The synchronizing

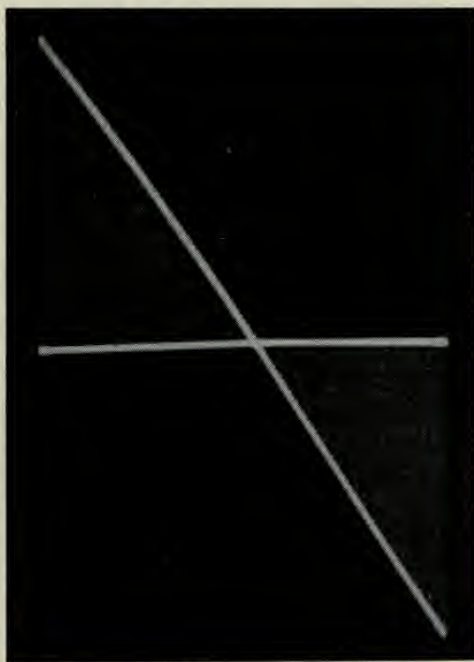


Fig. 24. Output of decoder (vertical) vs. input to coder (horizontal).

gear has been found equally easy to maintain. On the other side of the ledger, occasional breakdowns have pointed up the need for alarm and automatic replacement facilities in any version of the system which might be developed for commercial service.

A few measured characteristics are given in this section in addition to the compression and audio gain-frequency characteristics already shown (Figs. 14 and 11, respectively).

Input-Output Characteristics. The diagonal trace in the oscilloscope pattern of Fig. 24 shows the relationship between the input of a coder (horizontal deflection) and the output of the corresponding decoder (vertical). For this pattern a full-load audio signal was applied to the input of the odd

coder only (without passing through the compressor), and the output was taken directly from the common output of the two decoders. Thus the six odd channels took turns transmitting the signal while the six even channels produced the horizontal center-line. Uniform quantizing steps may be seen along most of the trace, but are obscured near the ends by defocusing of the test oscilloscope.

A similar pattern, obtained with the compressor and expander included in the transmission path, is shown in Fig. 25. Here tapered steps may be discerned, as well as a small amount of non-linearity due to residual imper-

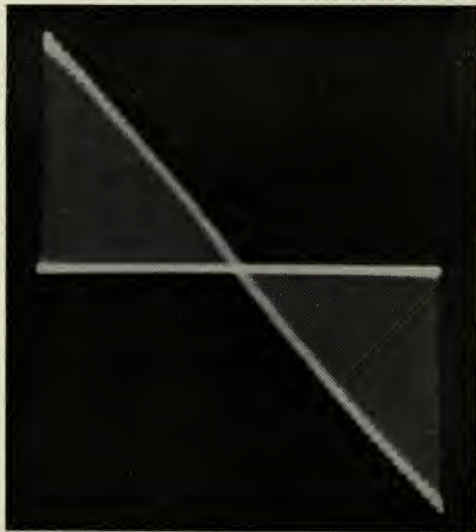


Fig. 25. Output of expander (vertical) vs. input to compressor (horizontal).

fections in the companding. The effect of the channel peak choppers is not included in this pattern.

Two measured overall input-output characteristics appear in Fig. 26, for the case of a typical single channel and for five channels patched in tandem through 17-decibel pads on a 4-wire basis. The latter simulates a possible extreme case of a long circuit in which for some reason it is desired to decode to audio at four junction points as well as at the final terminal. It should not be confused with a series of spans between which the PCM pulses are amplified or regenerated without decoding. In the latter case, of course, the overall audio characteristics are independent of the number of spans.

Audio Noise. Quantizing was found to be the only significant source of noise in the received audio signals. Noise levels measured in the absence of speech are shown in Fig. 27. The measurements are given for various num-

bers of channels from one to ten, connected in tandem as described in the preceding paragraph. Two scales of ordinates are shown in this figure. On

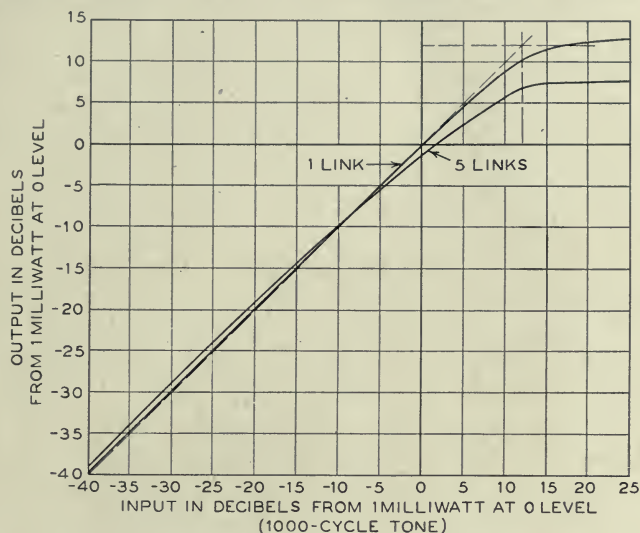


Fig. 26. Input-output characteristics of PCM channels.

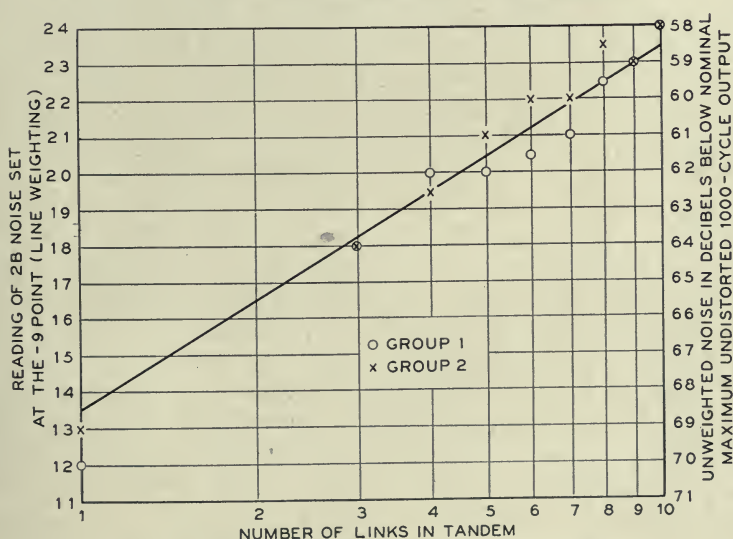


Fig. 27. Noise measurements on idle PCM channels.

the left is a reference scale of weighted noise employed in the Western Electric 2B Noise Set; on the right, a scale which relates the corresponding unweighted root-mean-square noise to the nominal maximum undistorted

single-frequency output of the system (taken to be 9 decibels above a milliwatt at zero level). This output corresponds to an input just reaching the peak limiters. Thus in a single link the idle circuit noise is down about 68 decibels from the full-load sine wave. For five links in tandem, the readings are at least 8 decibels below the accepted limit (29 decibels on the left-hand scale) for noise at the end of a 4000-mile circuit. Quantizing noise is found to increase approximately three decibels for each doubling of the number of links, as is generally the case with other forms of random noise.

For a single channel, or a small number of channels in tandem, the idle circuit noise varies considerably with the vertical centering of the coding tube. This may be understood by noting that if the zero-signal operating point effectively rests near the center of a "tread" on the quantizing staircase, (Fig. 4) a small amount of power hum or other disturbance may simply move it back and forth on the same tread, in which case quantizing noise is entirely absent. On the other hand, if the operating point is near a "riser" the small disturbances may cause it to joggle from one tread to another, producing noise. The measurements given in Fig. 27 were obtained with a very quiet input circuit and with the centering adjusted for maximum noise. The "joggling" was produced largely by residual power hum.

A-B tests to compare PCM transmission over a single link with direct transmission over a noise-free circuit of the same audio band were carried out using a wide range of talker volumes. In such tests only a few experienced observers were able to pick the PCM path consistently. When the PCM circuit included five tandem links most observers could tell the difference, but all judged the quality to be more than satisfactory for toll service.

Crosstalk. It has been pointed out earlier in this paper that interchannel crosstalk in a PCM system can occur only in the terminal equipment. Considerable care was exercised, particularly in the design of the time-division parts of the system, to hold individual sources of crosstalk to 70 decibels or better. As a result, measurements using a single-frequency test tone and a current analyzer have shown the overall crosstalk from any one channel to any other to be down 66 decibels in the worst cases.

Very severe tests have also been made in which a loud talker was connected to ten channels of a group simultaneously and crosstalk into either of the two remaining channels (one odd, one even) was listened to, and measured with the 2-B Noise Set. In such tests unintelligible crosstalk could be detected, which seemed to consist of changes in the quality of the quantizing noise occurring at a syllabic rate. The 2-B readings averaged about a decibel above the quantizing noise of a single idle channel with occasional peaks reaching the 17-decibel point on the reference scale.

In tests involving 24 talkers in 12 simultaneous conversations, crosstalk was practically undetectable.

Radio Interference and Noise. To obtain experimental confirmation of the expected tolerance to high interference levels in the radio path, the output of an oscillator, tunable through the band near 65 megacycles, was superimposed upon the received intermediate-frequency signal at the input to the group selection filters. With this controllable interference tuned near the center of the Group 1 filter, and its amplitude set 6 decibels below the peak amplitude of the (noise-free) pulses, errors were so plentiful that the demodulator did not remain synchronized. Proper framing was restored when the amplitude difference was increased to 7 decibels, but enough decoding errors remained to give intolerable audio noise. At 8 decibels only an occasional crackle of noise was observed, and at 9 decibels reception was perfectly normal. Similar tests of Group 2 gave the same results except, of course, that synchronization was not affected. The fact that noise-free transmission was not maintained quite up to the ideal 6-decibel point is due principally to the width (0.4 microsecond) of the gate which is applied to the rounded PCM pulses entering the receiving equipment. If necessary the ideal could undoubtedly be approached more closely by reducing this width, thus admitting only the extreme peaks and troughs of the signal.

The effects of actual fluctuation noise were studied by sending the PCM pulses over the radio path at reduced level. The boundary between good and bad transmission was not so sharp as with the continuous-wave interference, as should be expected because of the random nature of the noise. Flawless reception occurred when the root-mean-square signal at the peak of a pulse was greater by 18 decibels than the root-mean-square noise, both measurements being made at the output of a group selection filter.

VI. ACKNOWLEDGMENT

The system described is the result of the co-ordinated efforts of many people. In particular, the writers wish to express their appreciation to Mr. R. W. Sears for his development of the coding tube, and to acknowledge the contributions of certain others who have been actively concerned with various phases of the project. These include Messrs. R. L. Carbrey, A. E. Johanson, J. M. Manley, G. W. Pentico, A. J. Rack, P. A. Reiling, L. R. Wrathall, and the mechanics and wiremen who have kept to their usual high standards throughout countless circuit changes.

The project was under the overall direction of Mr. C. B. Feldman.

Electron Beam Deflection Tube for Pulse Code Modulation

By R. W. SEARS

INTRODUCTION

PULSE code transmission systems¹ in which successive signal amplitude samples are transmitted by pulse code groups require special modulators. The essential operational requirements of a pulse code modulator are: (1) to quantize or measure the signal amplitude sample to the nearest step in the discrete amplitude scale transmitted by the pulse code system, and (2) to generate the group of on-off pulses identifying the step.

Several methods have been proposed^{2, 3, 4} in which quantization and pulse formation were performed with circuits employing conventional electron tubes. The circuits involved sequential and comparison operations and were not easily adapted to a multi-channel time division system because of limitations in coding speeds and the complexity of the equipment. An electron beam deflection tube has been developed which, together with associated beam positioning and sweep circuits, performs the modulation rapidly, making possible the sequential modulation of a number of channels in time division multiplex.

The electronic principles, design and characteristics of the experimental tube are described in the present paper.

CONVERSION FROM SIGNAL INPUT TO PULSE CODE OUTPUT

An electrical input voltage may be converted into an output code pulse group with the electron beam deflection tube shown in Fig. 1a. An aperture or code masking plate is arranged perpendicular to the axis of the electron gun at the focal point. The coordinates of the aperture plate are aligned with the deflection axes of the X and Y deflector plate pairs. The electron beam strikes the output plate when it passes through an opening in the aperture plate.

An input voltage of appropriate value applied to the Y deflector plates will deflect the beam to point "a" of the aperture plate as indicated in Fig. 1a. A linear sweep voltage applied to the X deflection plates, while the

¹ An Experimental Multi-Channel Pulse Code Modulation System of Toll Quality, L. A. Meacham and E. Peterson, this issue.

² A. H. Reeves, *U. S. Patent* #2,272,070, Feb. 3, 1942.

³ H. S. Black and J. O. Edson, paper presented June 11, 1947 at A. I. E. E. meeting; Montreal, Canada.

⁴ W. M. Goodall, *Bell System Technical Journal*, July 1947.

input voltage on the Y deflection plates is held constant, causes the electron beam to sweep across the aperture plate along the dashed line $a-b$. A time sequence of output pulses is produced at the output plate when the beam passes through the apertures of the code plate along the path $a-b$.

A series of output pulses or a "pulse group" is characterized by the presence of pulses at time positions corresponding to the several vertical columns of apertures. The code plate shown in Fig. 1 is laid out in accordance with

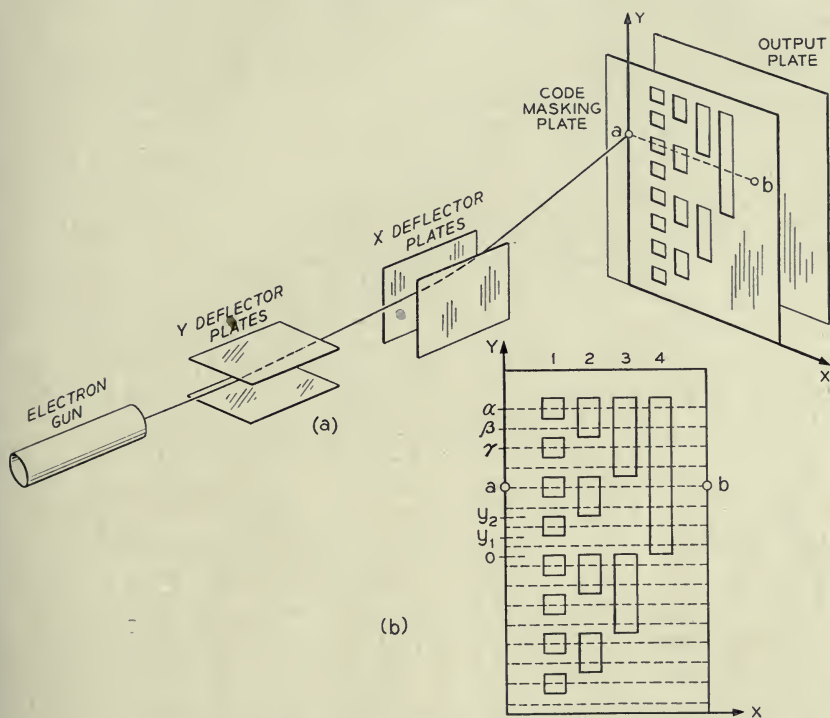


Fig. 1—Electron beam deflection tube for coding.

the binary number system pattern⁵ which was chosen for the present pulse code work because of the simplicity of decoding. The four-digit code plate shown in the tube of Fig. 1 provides for coding only 16 amplitude values and was used to facilitate the illustration. The tube developed for the ex-

⁵ The four vertical columns of the four-digit code plate in Fig. 1b provide four positions for output pulses. Vertical columns marked 1, 2, 3 and 4 correspond to the first, second, third and fourth digits, respectively, of the pulse code transmitted. Pulses located in time in accordance with this notation are given "weights" of 1, 2, 4 and 8, respectively. Sixteen pulse group combinations are indicated by the 16 dashed horizontal lines in Fig. 1b. The pulse group defined by beam sweep α corresponds to a total "weight" of 15; beam positions β and γ correspond to total "weights" of 14 and 13, respectively, and so on, with the bottom horizontal dashed line in the figure corresponding to zero.

perimental pulse code system uses a seven-digit code plate which provides for coding 128 amplitude values.

The vertical distances between the centers of successive code sweep positions (horizontal dashed lines in Fig. 1b) are made equal so that the codes are spaced by equal input voltage steps. A continuous range of input signal amplitudes will result in a continuous range of horizontal sweep positions. With an infinitely small electron beam, input signal amplitudes in the range from 0 to y_1 will produce a single output pulse group and input signal sample amplitudes from y_1 to y_2 will produce another output pulse group. This process of dividing the total input amplitude range into finite steps and arranging that input voltages falling within each step produce one and only one output pulse group is called quantization.

The tube of Fig. 1 will only quantize effectively if the electron beam is infinitely small and the sweep and aperture plate axes are aligned exactly. With a finite beam size, there will be sweep positions for which the beam straddles and sweeps out a combination of two adjacent codes.

Precise quantization and a uniform pulse output are required. The problems of quantizing, alignment and uniform pulse output have been solved by the use of a wire grid, called the quantizing grid, located in front of the aperture plate.

QUANTIZATION OF BEAM POSITION BY FEEDBACK

The quantizing grid consists of a horizontal array of grid wires aligned parallel to the code sweep or X axis of the aperture plate. The grid spacings and alignment are such that a wire lies between each adjacent pair of code groups as viewed from the various incident angles of the deflected electron beam. The quantizing grid, by means of an electrical feedback path to the signal deflection plates, divides the input signal range into a number of equal steps and positions the electron beam to the proper level for the code corresponding to the voltage step within which the signal amplitude sample falls. The quantizing grid wires also constrain the electron beam during the formation of the output code pulses, so that it must sweep out the code initially selected. In general, wires or shaped electrodes of any sort located where the electron beam can impinge thereon and connected in feedback relation to the deflection system constrains the electron beam to move in patterns prescribed by these electrodes and are thus called beam guides.

The coding tube with quantizing grid and feedback circuit is shown schematically in Fig. 2. The electrode line-up, reading from left to right in the figure, consists of an electron gun, deflection system, secondary electron collector, quantizing grid, aperture plate and output plate.

For the present purpose, a consideration of the collector and output plate electrodes is omitted and it is assumed that the grid does not emit secondary

electrons when bombarded. Electrons which strike the grid produce a current in the grid circuit while electrons that miss the grid have no effect in the grid circuit. The electron beam current intercepted by the grid will be dependent on the y deflection or position of the electron beam. The current is a maximum when the beam is centered on a grid wire and a minimum when it is centered between two grid wires and varies with beam position as indicated by the curve in the lower portion of Fig. 3. The curve is constructed for the case in which the beam diameter is slightly greater than the space between two grid wires. The current to the grid never becomes zero for any beam position. It may be thought of as having a "d-c. component" B . Amplifier 2 in Fig. 2 introduces a bias which cancels this "d-c. component" so that the feedback voltage is symmetrical about zero.

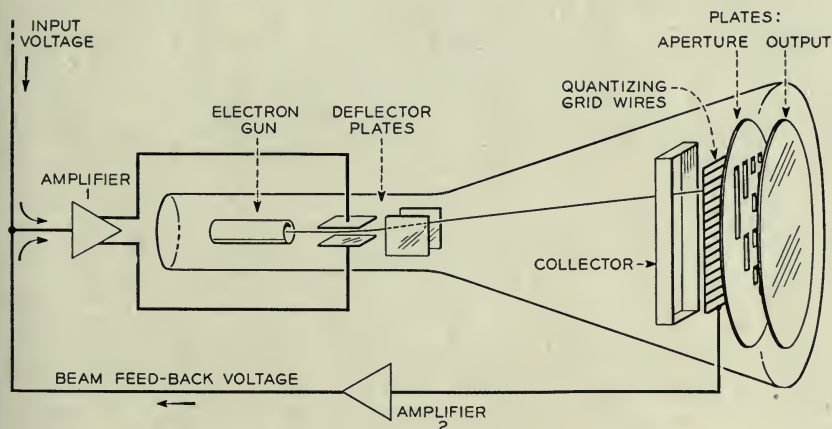


Fig. 2—Coding tube with quantizing grid and circuit schematic.

The grid current is amplified and the voltage developed is applied in feedback relation to the Y deflection plates as shown in Fig. 2. The case in which a positive feedback potential deflects the beam in the same direction as that for a positive signal voltage will be considered.

The beam deflecting voltage is equal to the sum of signal and feedback voltages and the beam position is a linear function of the deflection voltage so we may write

$$-e + Dy = e_f \quad (1)$$

where e is the input signal voltage, D the deflection constant, y the beam deflection or position and e_f the feedback voltage. The feedback voltage e_f is a periodic function of beam position y . Equation 1 therefore defines equilibrium beam positions for input signal voltage e . Equilibrium beam positions in accordance with Equation 1 are determined graphically in the

top portion of Fig. 3. The feedback voltage representing the right-hand side of Equation 1 is plotted as a function of beam position. The electron beam will have several possible positions of equilibrium at points (p_1 , p_2 , p_3 , p_4 , p_5 , p_6 and p_7) where the deflection line D erected from $-e$, representing the left-hand portion of Equation 1, crosses the feedback curve. These are the only beam positions for which the deflection potential (signal plus feedback) attains correct values for corresponding beam positions.

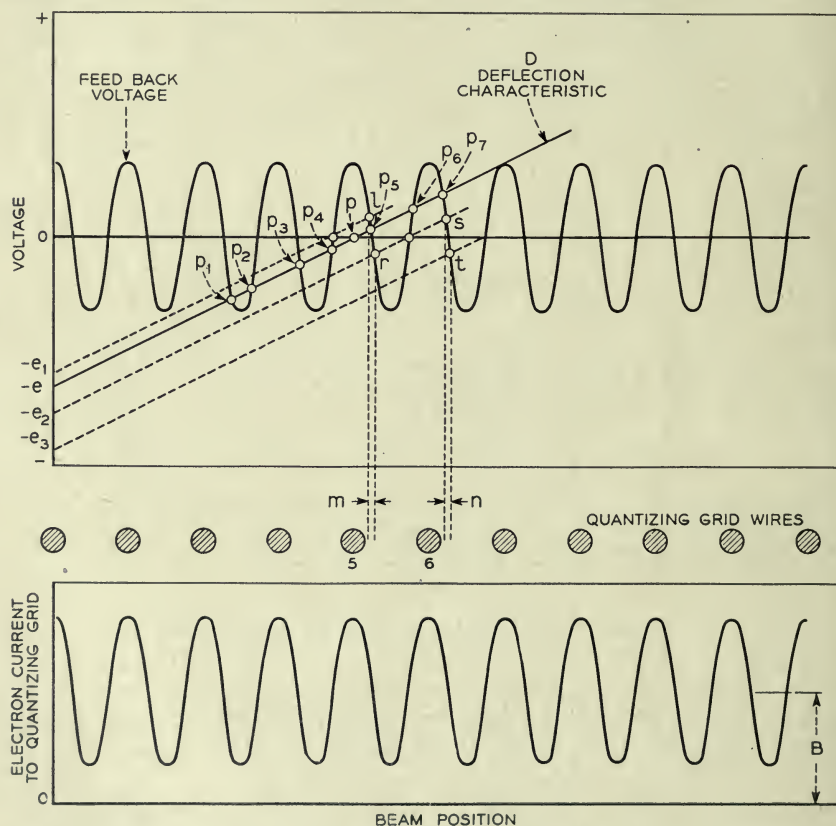


Fig. 3—Graphical representation of quantization.

However, only positions p_1 , p_3 , p_5 and p_7 will be in a true state of equilibrium. This can be seen as follows: Consider the beam at point p_3 ; if the beam is perturbed toward the right, the feedback voltage tending to deflect the beam to the left increases. The opposite action ensues if the beam is perturbed in the left-hand direction. Point p_3 is a true equilibrium point. On the other hand, consider the beam at position p_4 . If the beam is perturbed to the right, the feedback voltage tending to deflect it to the right increases and the beam continues to move until it reaches the equilibrium

position p_5 . If the beam is perturbed to the left from position p_4 it will continue to move to the left until it reaches position p_3 .

The number of possible equilibrium beam positions for an input signal sample depends on the maximum values of the feedback voltage and the slope of the deflection characteristic. It is necessary that only one equilibrium beam position be available for a particular small range of signal voltage. This can be achieved as follows: if a signal voltage e is established with the feedback circuit inoperative, the beam will be at a position p , Fig. 3. When the feedback circuit is activated with the signal voltage held at e , the beam will move from p to p_5 . With this procedure, signals in the range from e_1 to e_2 will result in equilibrium beam positions between points l and r on the curve. Thus, for signal voltages within the range from e_1 to e_2 the beam will fall in the small spacial interval m , whereas beam positions for input signals in the same range without feedback would vary from grid wire 5 to grid wire 6. Likewise, signal voltages between e_2 and e_3 will cause the beam to assume positions between s and t in the spacial interval n . The electron beam may be thought of as "leaning" on one side of a grid wire for a finite signal voltage range and on the same side of the next grid wire for an adjacent signal voltage range.

If the feedback voltage is of the opposite polarity to that assumed above, the quantizing action proceeds in the same manner except that the quantized beam positions lie at the left of the wires. The beam may be thought of as "leaning" on the opposite side of the grid wire.

The proper quantizing action is obtained by establishing and holding the signal voltage with the feedback circuit inoperative and then activating this circuit. The feedback circuit may be deactivated and activated by either (a) blanking and deblanking the electron beam, or (b) defocusing and focusing the electron beam by applying the proper voltage change to the beam control or focusing electrodes of the gun, respectively.

Since the grid wires are parallel to the horizontal rows of aperture holes, the feedback action constrains the beam to sweep out the code group initially selected even though the sweep axis is tilted slightly with respect to the grid wires and aperture plate. The maximum swing of the feedback voltage at the deflection plates should be about three or four times the value of the voltage required to deflect the beam from one code group to the next in order to provide ample protection against the beam jumping from one code group to the next code group during the sweep.

THE EXPERIMENTAL CODING TUBE

The seven-digit experimental tube developed for pulse code transmission system trials utilizing the electrode lineup shown schematically in Fig. 2 is pictured in Fig. 4. The electron gun assembly and the target plate assembly

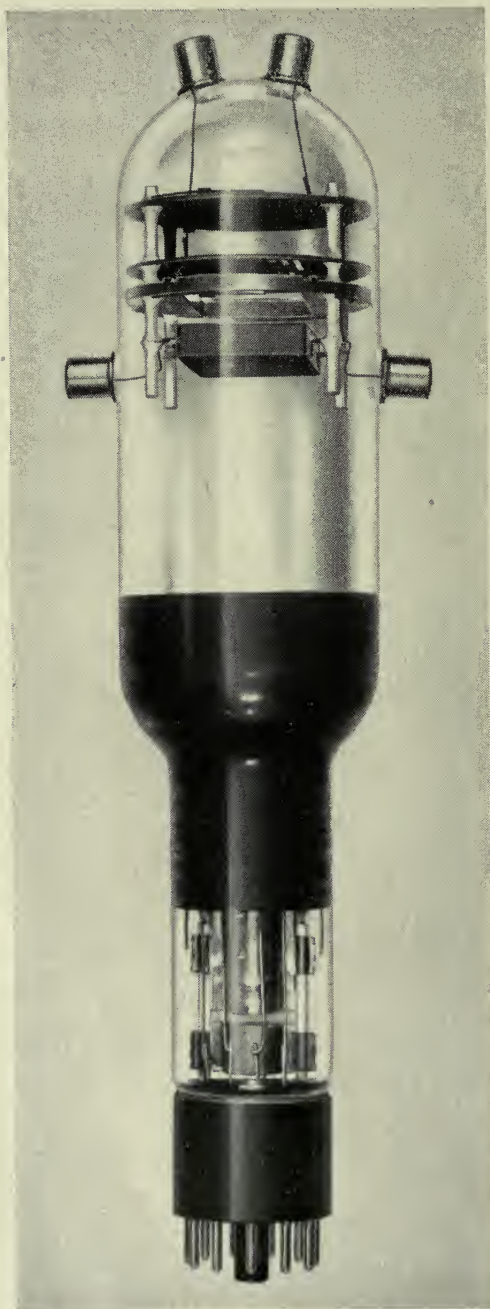


Fig. 4—Experimental seven-digit coding tube.

are sealed in at opposite ends of the tube envelope. The over-all length of the tube is $11\frac{1}{4}$ " with a maximum bulb diameter of $2\frac{1}{4}$ ".

The electron gun operates at a final anode potential of 1000 volts with a beam current of approximately 10 microamperes. A potential of about 100 volts applied to the signal input deflection plates deflects the beam from the center to the top of the aperture plate. This corresponds to a maximum deflection angle of $10\frac{1}{2}^\circ$.

The four electrodes of the target assembly, secondary collector, quantizing grid, aperture plate and output plate are shown in the photograph of Fig. 5 from bottom to top respectively. The secondary collector is a simple rectangular shaped electrode. The quantizing grid consists of a circular frame with a parallel array of grid wires stretched across a rectangular opening. The aperture plate is a thin disc with apertures arranged in a binary pattern which provides for a seven-digit code. The output plate is a thin circular disc. Both aperture and output plates are coated with a carbon layer to suppress secondary electron emission from their surfaces.

The parts of the target assembly are held in accurate alignment in a jig and cemented and held in position on four ceramic rods. The entire assembly is held rigidly in the tube envelope by means of spacers attached to the quantizing grid and output plate.

The target assembly is aligned with the electron gun and deflection plate axes by means of lineup tools in the glass lathe at the time the final seal is made at the center of the glass envelope. It has been possible to hold the alignment of the deflection axes with the aperture plate to within slightly less than 1° with this construction.

The construction of the quantizing grid may be seen more clearly in the photograph of Fig. 6. The grid frame has raised portions on two sides of the rectangular opening. These are milled with a series of grooves for each grid wire. The grid laterals are affixed in the grooves by brazing and are thus accurately spaced with respect to each other and to assembly lineup holes which can be seen spaced around the edge of the grid frame. The wires are held taut by means of a flat spring which is welded to the grid frame and supplies tension to stretch the lateral wires. The grid wires are 4.0 mils in diameter, processed to have a secondary emission coefficient of about 3. The laterals are spaced 11.6 mils between centers.

The openings in the aperture plate are made by a punching operation. The area of the aperture plate covered by the first and second digit columns is milled to a thickness of 5 mils in order to facilitate accurate punching of the smallest apertures. The apertures in the first digit column (bottom horizontal row in Fig. 5) are rectangular .012" x .062". The seven-digit code pattern provides 128 different output pulse groups. The wide openings

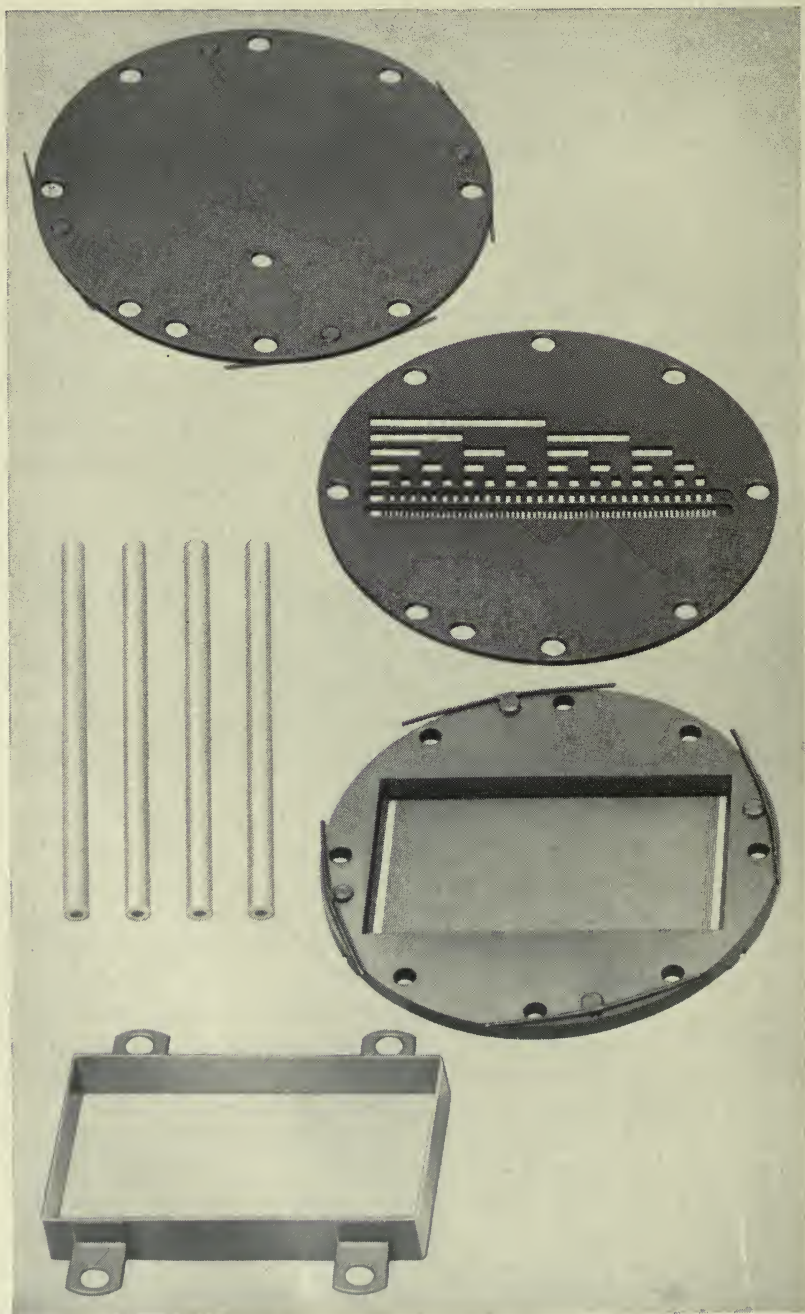


Fig. 5—Target electrodes.

at the left hand side of the aperture plate in Fig. 5 provide a peak amplitude range for which the output pulse group consists of all seven pulses. This in effect provides a peak limiting action.

Leads from the four electrodes of the target assembly are brought out directly to terminal caps on the side and end of the tube envelope to decrease the interelectrode capacitances and to facilitate direct connection to external circuits.

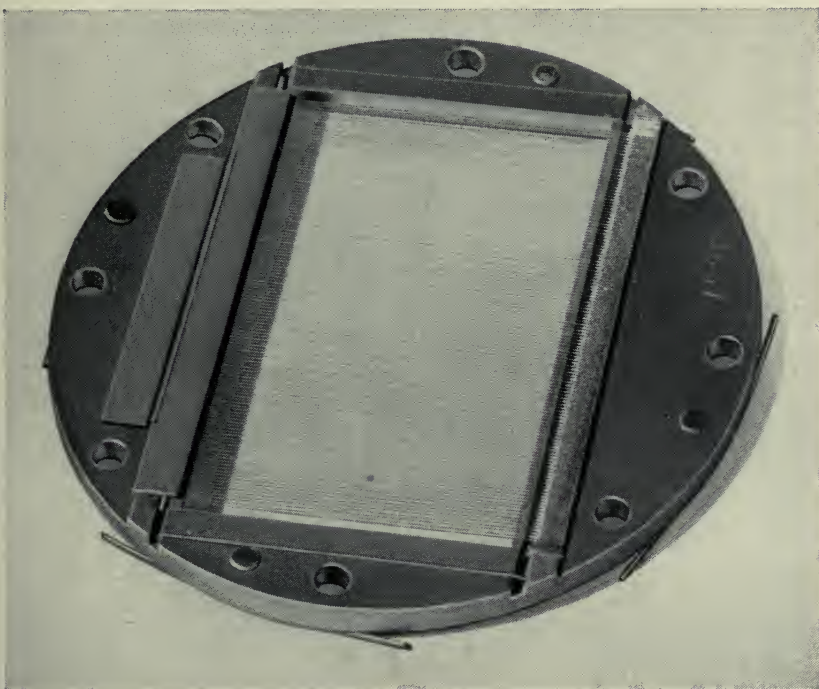


Fig. 6—Quantizing grid.

OPERATIONAL AND DESIGN CONSIDERATIONS

The quantizing action depends on the periodic variation of electron beam current intercepted by the wires of the quantizing grid with beam position. Secondary electron emission from the grid and other electrodes was assumed to be negligible with the feedback circuit connected directly to the quantizing grid as shown in Fig. 2. The uniform suppression of secondary electrons is difficult to achieve even though the grid is coated and processed for a low secondary ratio. It is also difficult to prevent secondary electrons produced at the aperture plate from being collected by the grid.

A preferred method of operation utilizes the secondary electrons produced

at the grid by the impinging primary beam for the quantizing action rather than the directly intercepted electron current as heretofore assumed. The secondary electron collector located in front of the quantizing grid is maintained at a positive potential and collects most of the secondaries from the grid. There is, of course, a correspondence between the secondary electron current and the fraction of the beam current intercepted by the grid wires. The quantizing circuit is made by connecting the feedback path to the secondary collector and the quantizing action proceeds as described heretofore. This method has the following advantages over the direct primary current method: (1) the collector current as a function of beam position is much more regular; (2) the swing between maximum and minimum current is considerably larger because of the secondary emission multiplication at the grid surface; and (3) the capacitance between collector and ground is lower than the capacitance between the closely spaced grid and aperture plate.

With secondary electron current feedback, the aperture plate is operated at a positive potential relative to the grid to suppress secondary electrons from the aperture plate. The proportion of the secondary emission from the grid collected by the aperture plate is small compared with that collected by the secondary collector. High velocity secondaries originating at the aperture plate are, however, able to penetrate the retarding field and strike the grid. These energetic secondaries produce low-velocity secondaries at the grid which flow to the secondary collector electrode. This alters the character of the secondary collector current in accordance with the spacial pattern of the apertures in the code plate. The surface of the aperture plate is carbonized to reduce the emission of high-velocity secondaries. The spacial variation of the quantizing current is reduced to less than 10% of the total quantizing current swing with this treatment.

The periodic nature of the quantizing current with beam position must be as uniform as possible as regards both the "a-c" and "d-c" components. The maximum swing should also be as large as possible to permit the use of the lowest possible impedance in the feedback path to obtain a wideband characteristic for fast coding.

The factors which determine the maximum current swing are electron beam current, secondary emission coefficient of the grid wires, electron beam spot size, grid wire diameter and spacing. The quantized beam falls approximately halfway between the center of a grid wire and the midpoint between wires. The beam must be small enough that its edge does not extend into the region beyond the wire on which it "leans" by an appreciable amount. This is obtained with the electron beam focused to a spot slightly smaller than that for maximum quantizing current amplitude.

Optimum performance has been obtained with the electron beam focused to a radius⁶ of about 5 mils for the grid spacings previously specified.

The principal irregularities in the periodic quantizing current are caused by variations in secondary emission coefficient of the grid and aperture plate and deflection defocusing of the electron beam. The secondary emission from the grid surface can be made sufficiently uniform by careful processing. This is illustrated in Fig. 7 which is a trace of an oscilloscope



Fig. 7—Variation of quantizing current with beam deflection.

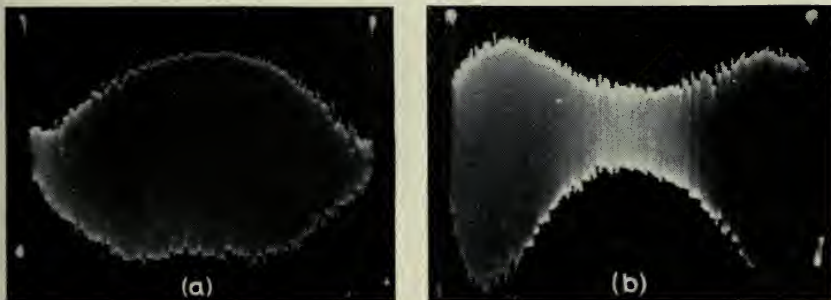


Fig. 8—Deflection defocussing.

presentation of the secondary collector current as the beam is deflected by a linear sweep at right angles to the grid laterals. The trace was limited to cover a sufficiently small number of wires to show details of its shape. Figure 8a illustrates the effect of deflection defocusing. The sweep was expanded to cover all of the grid wires and the electron beam focused for maximum amplitude in the center region. The curve is so compressed that

⁶ Distance from the center of the beam to a point at which the electron current density is 5% of its value at the center.

the individual oscillations of the current between maximum and minimum values can hardly be resolved. The envelope of the curve indicates the extent to which the swing is reduced at the two ends by the increase in electron beam spot size which results from deflection defocusing. In Fig. 8b, the focusing voltage has been changed by 12 volts and it can be seen that the beam is in focus at the maximum deflection angles and out of focus in the center region. With an intermediate or compromise focus voltage, the tube will quantize satisfactorily over the entire code range. Best results have been obtained with the tube in the experimental pulse code system by the use of a simple circuit which changes the focus voltage in a linear manner with the rectified or absolute value of the input deflecting signal thus compensating for the deflection defocusing of the tube.

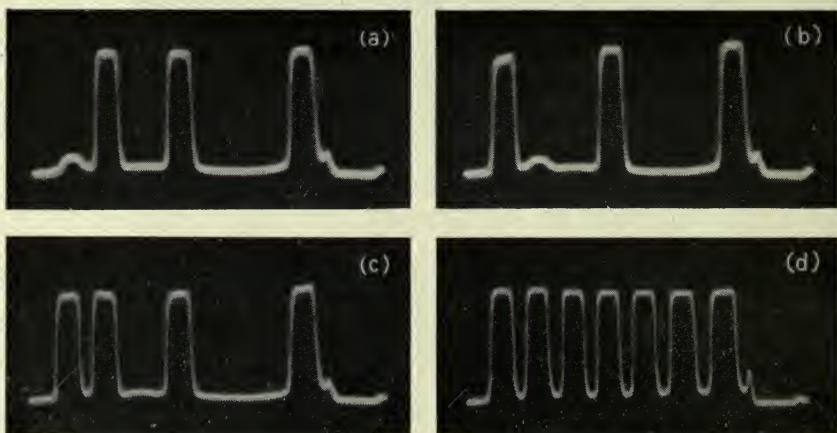


Fig. 9—Typical pulse code outputs.

The output plate is usually operated at a positive potential relative to the aperture plate with attendant suppression of secondary emission from the former. Output pulses of opposite polarity may be obtained by operating the output plate at a negative potential and processing this plate to have a secondary emission ratio greater than unity.

The several groups of output pulses shown in Fig. 9 are illustrative of the tube output. Successive digit pulse positions occur from left to right in the figures. The small kink in the base line at the first and second digit pulse positions in codes a and b respectively are small error pulses caused by a small portion of the edge of the beam overlapping a grid wire when the beam is quantized. Error pulses of this magnitude are readily eliminated by slicing when the tube is used in the pulse code transmission system.

Coding tubes have operated satisfactorily for long periods of time in the experimental multi-channel pulse code system and have required minor adjustments of potentials no more than once a day. The tube operates in this equipment with a code sweep time of 10 microseconds.

ACKNOWLEDGMENT

The writer wishes to acknowledge the original suggestions of Dr. F. B. Llewellyn and Mr. G. Hecht and the help of his colleagues. The contributions and wholehearted cooperation of Messrs. C. B. Feldman, L. A. Meacham and their associates who concurrently developed the pulse code system and circuits associated with the coding tube, have been invaluable. Mr. A. Salecker deserves a great deal of credit for the detailed mechanical design of the tube and his work in producing tube models.

The development was under the over-all direction of Mr. J. R. Wilson and Dr. S. B. Ingram.

Metallic Delay Lenses

By WINSTON E. KOCK

A metallic lens antenna is described in which the focussing action is obtained by a reduction of the phase velocity of radio waves passing through the lens rather than by increasing it as in the original metal plate lens. The lens shape accordingly corresponds to that of a glass optical lens, being thick at the center and thin at the edges. The reduced velocity or "delay" is caused by the presence of conducting elements whose length in the direction of the electric vector of the impressed field is small compared to the wavelength; these act as small dipoles similar to the molecular dipoles set up in non-polar dielectrics by an impressed field. The lens possesses the relatively broad band characteristics of a solid dielectric lens, and since the conducting element can be made quite light, the weight advantage of the metal lens is retained. Various types of lenses are described and a theoretical discussion of the expected dielectric constants is given. An antenna design which is especially suitable for microwave repeater application is described in some detail.

INTRODUCTION

THE metal lens antennas described by the writer elsewhere¹ comprised rows of conducting plates which acted as wave guides; a focussing effect was achieved by virtue of the higher phase velocity of electromagnetic waves passing between the plates. Higher phase velocity connotes an effective index of refraction less than unity, and a converging lens therefore assumes a concave shape. The relation between the index of refraction n , the plate spacing, a , and the wavelength λ

$$n = \sqrt{1 - (\lambda/2a)^2}, \quad (1)$$

indicates that the refractive index varies with wave length. As a consequence, such lenses exhibit "chromatic aberration"; i.e. the band of frequencies over which they will satisfactorily operate is limited. Although some of these lenses may have ample bandwidth (15% to 20%) for most microwave applications, others, having large apertures in wavelengths, may have objectionable bandwidth limitations. For example, the lens of Fig. 1, having an aperture diameter of 96 wavelengths, has a useful bandwidth of only 5%.

As a means for overcoming these band limitations, the metallic lenses of this paper were developed. They are light in weight and possess an index of refraction which can be made sensibly constant over any desired band of microwave frequencies. They avoid the weight disadvantages of glass or plastic lenses, and retain the tolerance and shielding advantages of the lens

¹ W. E. Kock, *Bell Laboratories Record*, May 1946, p. 193; *Proc. I. R. E.*, Nov. 1946, p. 828.

over the reflector antenna. Because electromagnetic waves passing through them are slowed down or "delayed" (as in the glass lenses of optics), they are called delay lenses; and since the elements in the lenses which produce the delay are purely metallic, they are called metallic delay lenses.

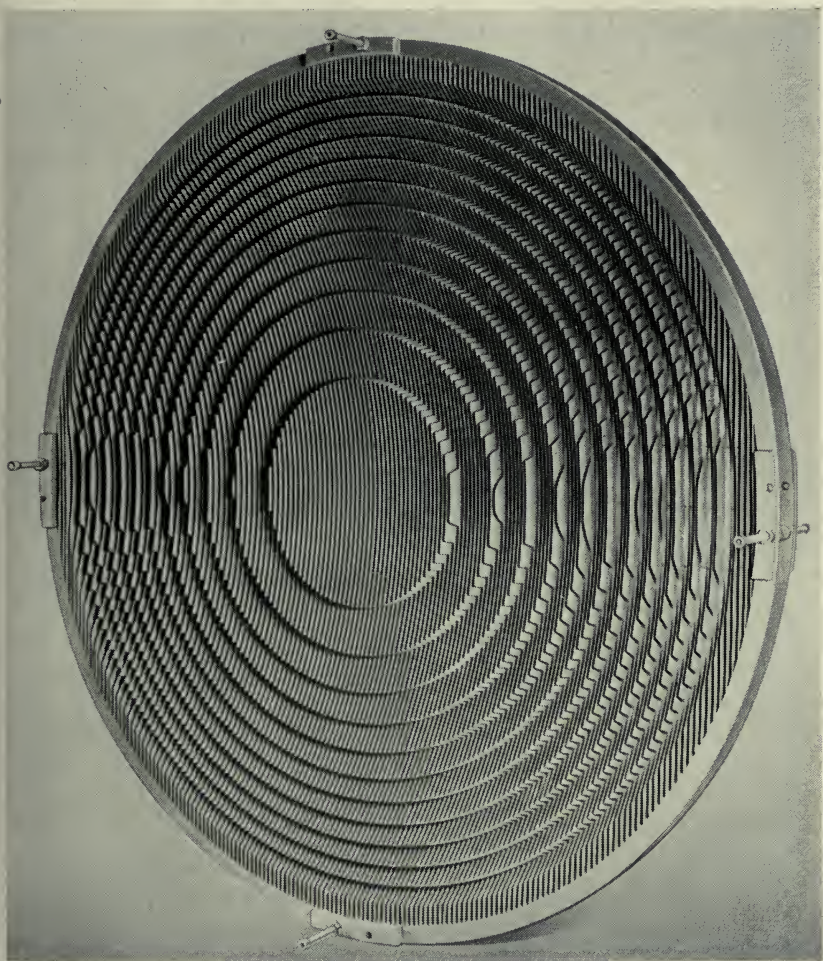


Fig. 1—Waveguide metallic lens having an aperture of 96 wavelengths and a useful bandwidth of 5%.

PART I—EXPERIMENTAL

FUNDAMENTAL PRINCIPLES

The artificial dielectric material which constitutes the delay lens was arrived at by reproducing, on a much larger scale, those processes occurring

in the molecules of a true dielectric which produce the observed delay of electromagnetic waves in such dielectrics. This involved arranging metallic elements in a three-dimensional array or lattice structure to simulate the crystalline lattices of the dielectric material. Such an array responds to radio waves just as a molecular lattice responds to light waves; the free electrons in the metal elements flow back and forth under the action of the alternating electric field, causing the elements to become oscillating dipoles similar to the oscillating molecular dipoles of the dielectric. In both cases, the relation between the effective dielectric constant ϵ of the medium, the density of the elements N (number per unit volume) and the dipole strength (polarizability α of each element) is approximately given by

$$\epsilon = \epsilon_0 + N\alpha \quad (2)$$

where ϵ_0 is the dielectric constant of free space.

There are two requirements which are imposed on the lattice structure. First, the spacing of the elements must be somewhat less than one wavelength of the shortest radio wave length to be transmitted, otherwise diffraction effects will occur as in ordinary dielectrics when the wavelength is shorter than the lattice spacing (X-ray diffraction by crystalline substances). Secondly, the size of the elements must be small relative to the minimum wavelength so that resonance effects are avoided. The first resonance occurs when the element size is approximately one half wavelength, and for frequencies in the vicinity of this resonance frequency the polarizability α of the element is not independent of frequency. If the element size is made equal to or less than one quarter wavelength at the smallest operating wavelength, it is found that α and hence ϵ in equation 2 is substantially constant for all longer wavelengths.

Since lenses of this type will effect an equal amount of wave delay at all wavelengths which are long compared to the size and spacing of the objects, they can be designed to operate over any desired wavelength band. For large operating bandwidths, the stepping process² is to be avoided, since the step design is correct only at one particular wavelength. Such unstepped lenses are thicker, but the diffraction at the steps is eliminated and a somewhat higher gain and superior pattern compared to a stepped lens is achieved. By tilting the lens a small amount, energy reflected from it is prevented from entering the feed line and a good impedance match between the antenna and the line can be maintained over a large band of frequencies.

Another way of looking at the wave delay produced by lattices of small conductors is to consider them as capacitative elements which "load" free

² The lens of Fig. 1 has 12 steps.

space, just as parallel capacitors on a transmission line act as loading elements to reduce the wave velocity. Consider a charged parallel plate air condenser with its electric lines of force perpendicular to the plates. Its capacity can be increased either by the insertion of dielectric material or by the insertion of insulated conducting objects between the plates if the objects have some length in the direction of the electric lines of force. This is because such objects will cause a rearrangement of the lines of force (with a consequent increase in their number) similar to that produced by the shift, due to an applied field, of the oppositely charged particles comprising the molecules of the dielectric material. The conducting elements in the lens may thus be considered either as portions of individual condensers, or as objects which, under the action of the applied field, act as dipoles and produce a dielectric polarization, similar to that formed by the rearrangement of the charged particles comprising a non-polar dielectric.³ Either viewpoint leads to the delay mechanism observed in the focussing action of the artificial dielectric lenses to be described.

EXPERIMENTAL MODELS

We turn now to experimental exemplifications of lenses built in accordance with the principles outlined above.

(a) *Sphere Array*

One of the simpler shapes of conducting elements to be tried was the sphere. Figure 2 is a sketch of an array of conducting spheres arranged approximately in the shape of a convex lens. The spheres are mounted on insulated supporting rods; the microwave feed horn and receiver are shown at the right. The focal length is f , the radius of the lens "aperture" is y , the maximum thickness is x and not only the spacings s_1 and s_2 but also the size of the spheres are small compared to the wavelength. Rays A and B are of equal electrical length because ray A is slowed down or delayed in passing through the lens. Figure 3 is a photograph of the lens of Fig. 2; it also portrays a similar sphere array lens made of steel ball bearings supported by sheets of polystyrene foam⁴ which have holes drilled in them to accept the spheres. In both cases the balls are arranged in a symmetrical lattice. It will be shown below that the polarizability α of a conducting

³ Polar dielectrics have arrangements of charged particles which are electric dipoles even before an external electric field is applied; the field tends to align these and the amount of polarization (and hence the dielectric constant) that they exhibit depends upon temperature, since collisions tend to destroy the alignment. Non-polar (or hetero-polar) molecules have no dipole moment until an electric field is applied; the polarization of such materials (and of the artificial dielectrics we are discussing) is accordingly independent of temperature. See, for example, Debye, "Polar Molecules", Chap. III.

⁴ Styrofoam (Dow). Because of its low density (1 to 2 pounds per cu. ft.), its contribution to the wave delay is negligible ($\epsilon_r = 1.02$).

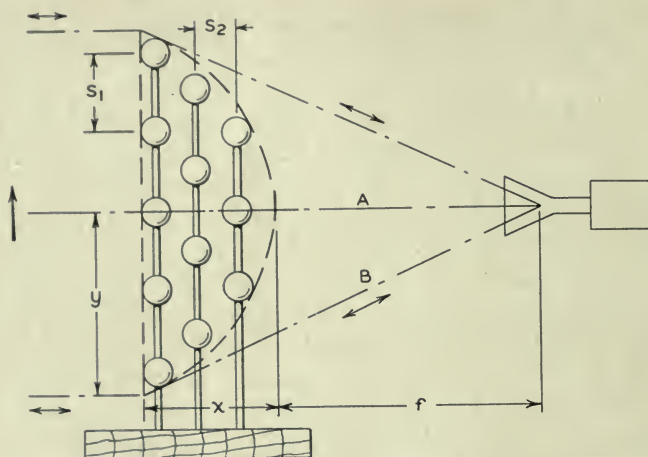


Fig. 2—Lattice of conducting spheres arranged to form a convex lens.

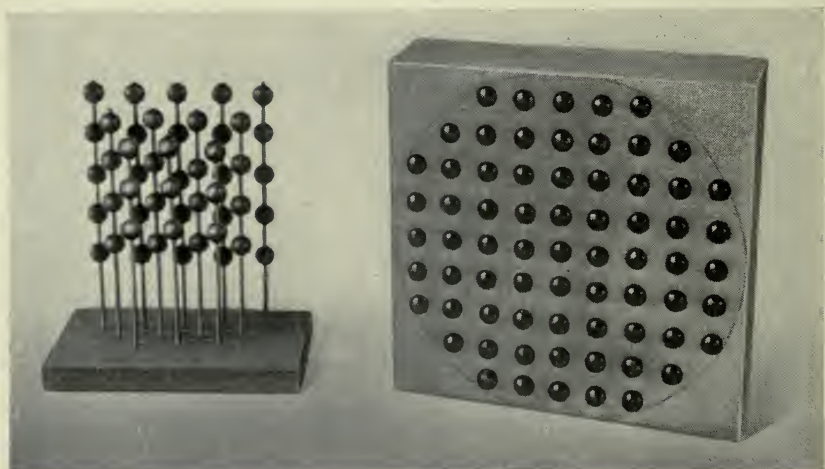


Fig. 3—Left: The sphere lens of Fig. 2. Right: Sphere lattice supported by foam sheets.

sphere of radius a is $4\pi\epsilon_0 a^3$, so that, for static fields, the relative dielectric constant is, from (2),

$$\epsilon_r = \epsilon/\epsilon_0 = 1 + 4\pi N a^3, \quad (3)$$

where N = number of spheres per unit volume.

For most dielectrics, the index of refraction is simply the square root of the relative dielectric constant. However, in the case of the sphere lens at microwaves eddy currents on the surface of the spheres prevent the magnetic

lines of forces from penetrating them and it will be seen later that this effect causes the expected value of the index of refraction to be smaller than that determined by the usual equation

$$n^2 = \epsilon_r. \quad (4)$$

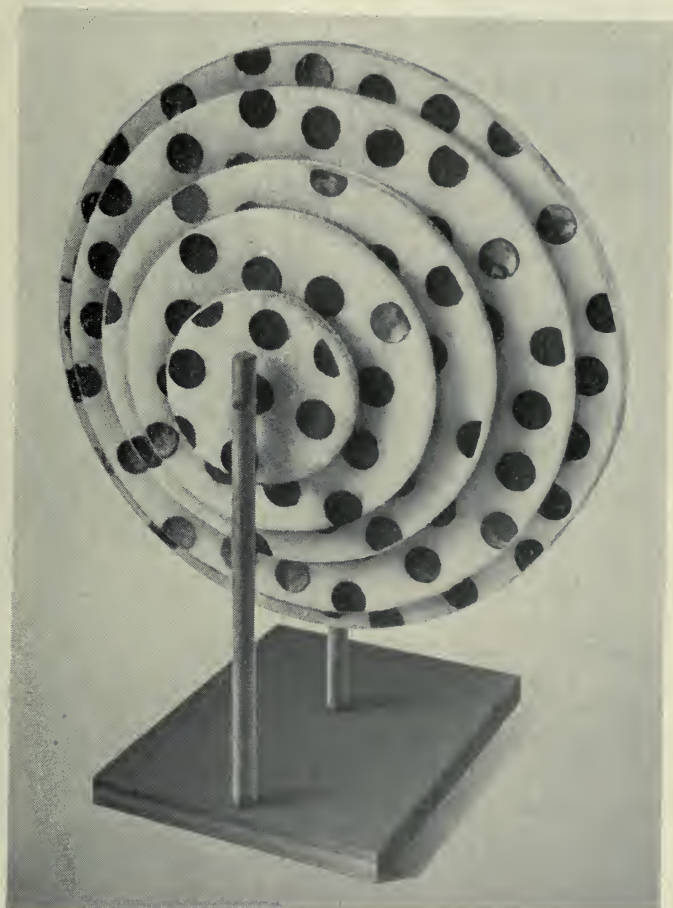


Fig. 4—Lattice of conducting disks arranged to form a plano-convex lens. Polystyrene foam sheets support the disks.

To avoid this effect the elements should be shaped so as not to alter the magnetic lines of force. This can be done by making their dimension in the direction of propagation of the impressed waves negligibly small.

(b) *Disk Array*

One accordingly arrives at the lens design shown in Fig. 4 in which the spheres are replaced by copper foil disks lying in planes parallel to the

impressed E and H vectors. The supporting sheets are again polystyrene foam. The foil disks have negligible thickness so that the magnetic lines are unaffected as shown in Fig. 5. Equation (4) is therefore valid even at radio frequencies and the index of refraction of this material is obtainable from (4) and (2) with α equal to $\frac{16 \epsilon_0 a^2}{3}$, as shown in the last section.

(c) *Strip Array*

Both the sphere and disk type lenses have the advantage that they will perform equally well on horizontally or vertically polarized waves. If the lens is required to focus only one type of wave polarization the disks can be replaced by thin, flat, conducting strips extending in length in the direction

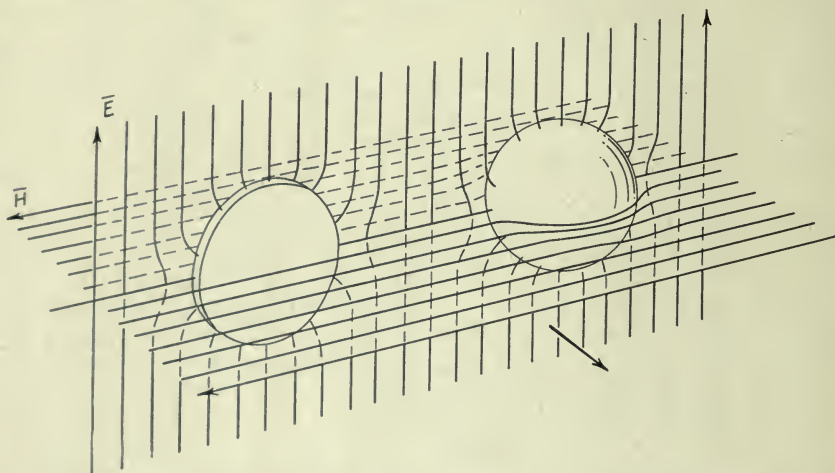


Fig. 5—Disturbance of the magnetic field is avoided by the use of disks instead of spheres.

of the magnetic vector of the applied field. A simple method of constructing such a lens is shown in Fig. 6. Slabs of dielectric foam are slotted to a depth equal to the strip width and each slab is marked with the profile contour necessary to produce the final plano-convex lens. The metal strips are then cut to the length indicated on the profile and inserted in the slots. With the strips in place the unit slabs are stacked on top of one another and held in a mounting frame to form the complete lens. Figure 7 shows one slab of a 10-foot strip type lens being assembled and Fig. 8 shows a six-foot square lens half assembled. Figure 9 shows directional patterns of a 3-foot diameter lens of this type made up of $\frac{3}{4}$ inch $\times$.005 inch strips spaced $\frac{3}{8}$ in the slabs and the slabs $1\frac{1}{2}$ inch thick. The three patterns were taken over a 12% band of frequencies with the lens purposely illuminated by

a low directivity feed at the focal point. This produced an almost uniform illumination across the aperture at all three wavelengths so that the side lobes were not too well suppressed. However, the deep minima in all three directional patterns indicate the relative absence of curvature of the emerging phase front; this signifies that the strip dielectric has a negligible variation of refractive index over the indicated wavelength band.

(d) *Sprayed Sheet Lenses*

The disk lens or strip lens can be constructed in the manner indicated in Fig. 10, which shows two lenses made by spraying conducting paint on thin dielectric sheets through masks. This results in round dot or square dot

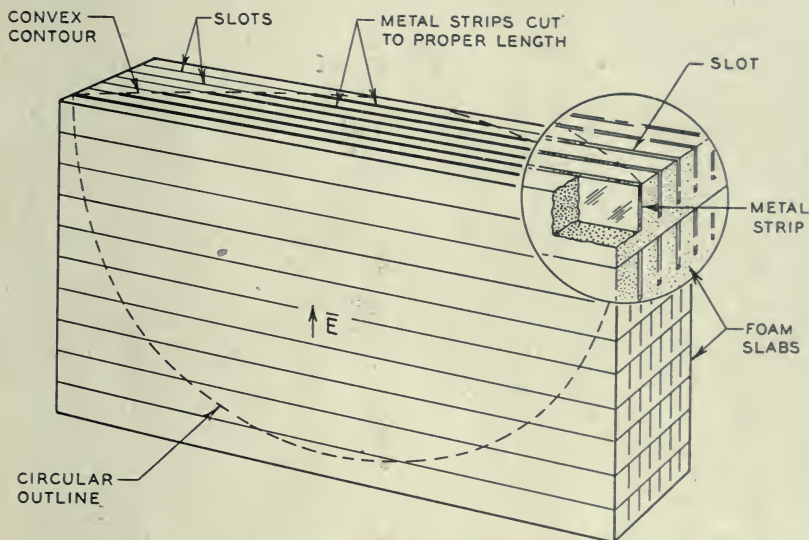


Fig. 6—Construction of a delay lens employing metallic strips as the delay elements.

patterns on the sheets and the size circle used on each sheet determines the three-dimensional contour when the sheets are stacked up to form the lens. Those in the photograph are spaced by wooden spacers as shown in the sketch of Fig. 11; for large lenses it would probably be preferable to cement the sheets to polystyrene foam spacers, thereby making a solid foam lens. Because of the small size of the elements, the lens on the left in Fig. 10, was effective at wavelengths as short as 1.25 cm, in addition to longer wavelengths (3, 7 and 10 cm). The lens of Fig. 12 was made by metal spraying metallic tin directly on circular foam slabs through masks, and the foam disks then cemented together.

Strip lenses can also be constructed in this way. The lens in Fig. 13,



Fig. 7—Inserting metal strips into a profile plate of a 10-foot strip lens.



Fig. 8—A six-foot square strip lens half assembled.

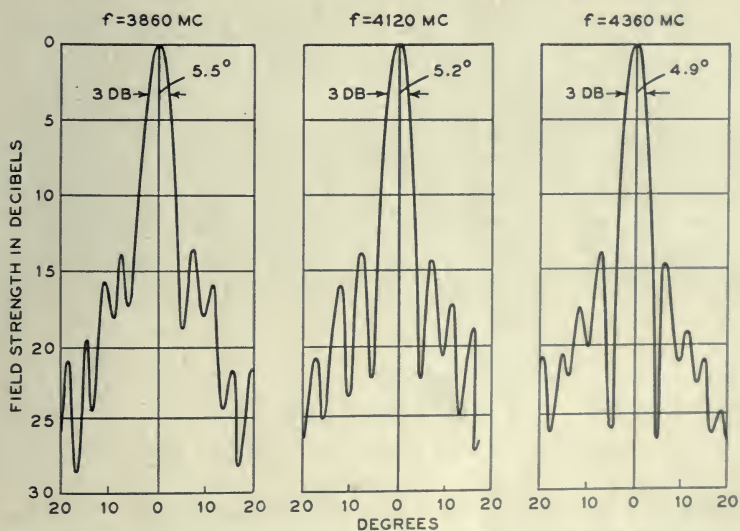


Fig. 9—Directional patterns of a 3-foot diameter lens at several frequencies.

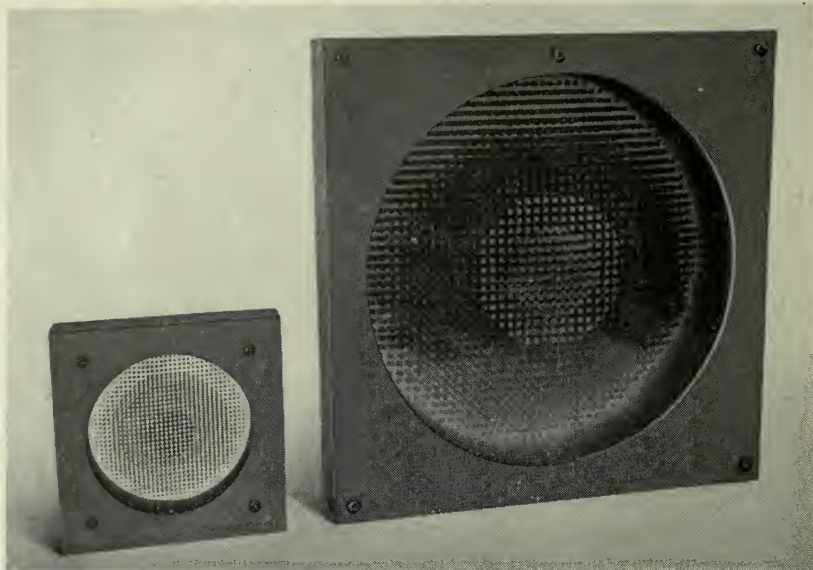


Fig. 10—Lenses constructed by spraying conducting paint on cellophane sheets.

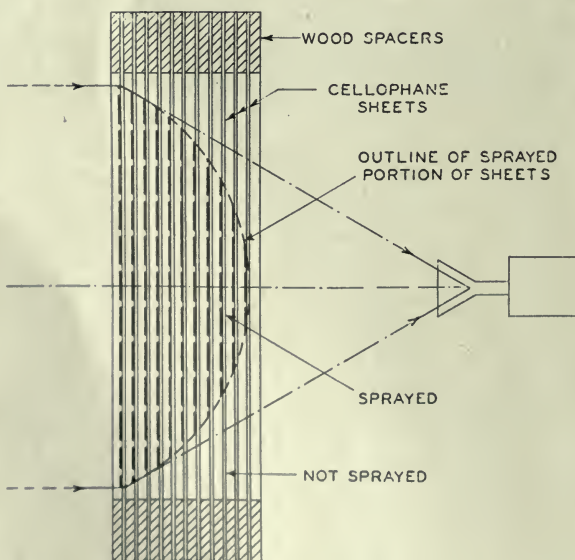


Fig. 11—Construction details of the lenses of Fig. 10.

was made by affixing copper foil strips to cellophane sheets and stacking the sheets with no spacing other than the sheets themselves between them.

This extremely close spacing and the staggered arrangement of strips correspond to a heavy capacitive loading and introduced so much delay per unit length that the measured effective dielectric constant of this lens proved to be 225 (i.e. the index of refraction was 15). Because of such a high dielectric constant the reflection losses at the surface of this lens are high,

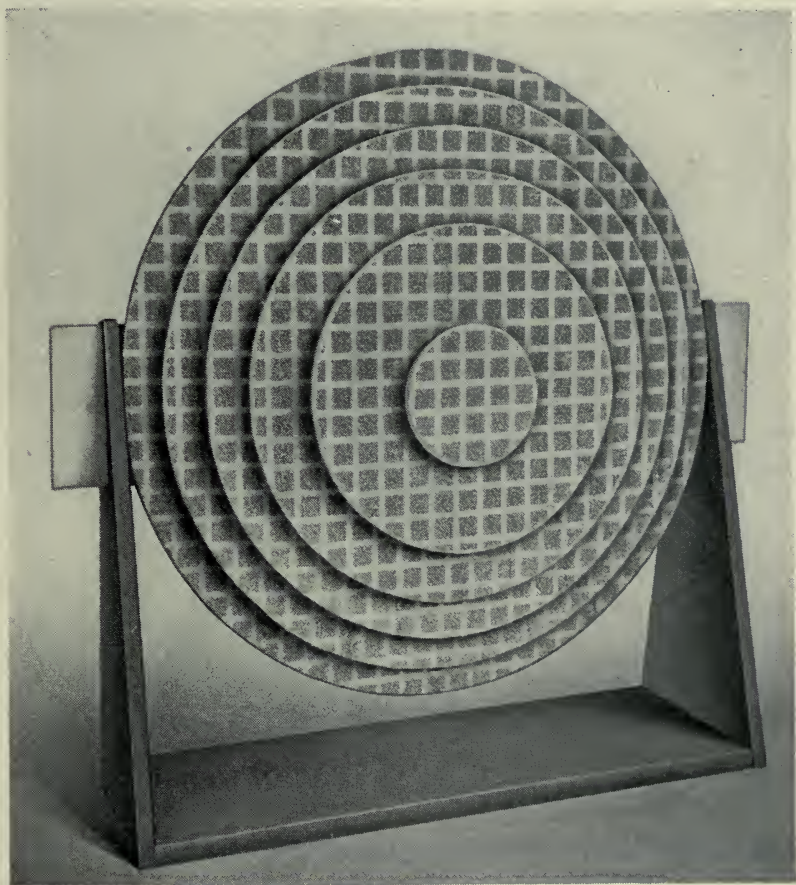


Fig. 12—Lens formed by metal-spraying tin directly onto polystyrene foam sheets.

so that, for efficient operation, tapered or quarter wave matching sections on each surface would be necessary.

(e) *Frequency Sensitive Delay Lenses*

When the conducting elements approach a half wavelength in their length parallel to the electric vector, resonance effects occur and the artificial dielectric behaves like an ordinary dielectric near its region of anomalous

dispersion.⁵ The index of refraction of an artificial dielectric using $\frac{3}{4}$ " metallic elements would increase rapidly as the wavelength approached $1\frac{1}{2}$ " until, at $\lambda = 1\frac{1}{2}$ " the dielectric would be opaque. At still lower wavelengths, the material would appear to have an index of refraction less than one.

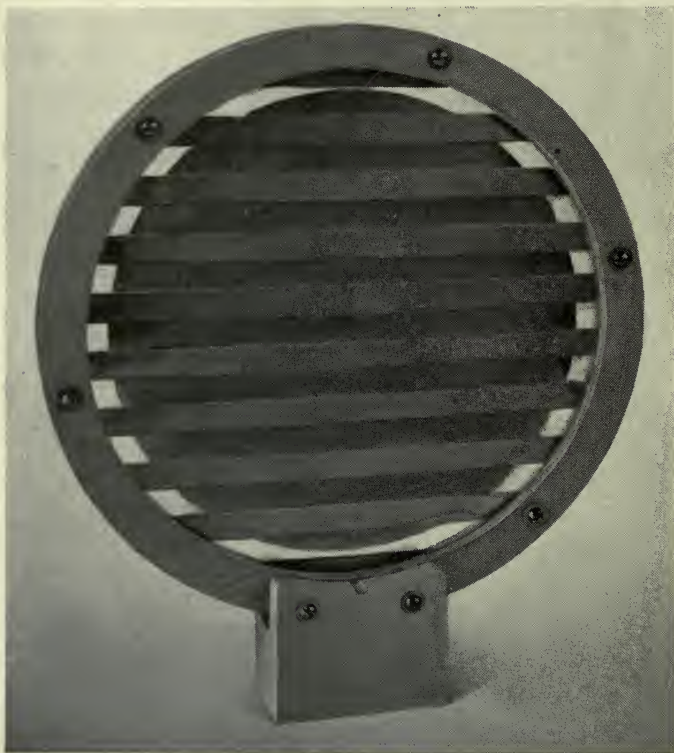


Fig. 13—Closely spaced strip construction comprising copper foil strips affixed to cellophane sheets. Juxtaposition of the sheets yielded an effective index of refraction of 15.

Elements, such as rods, which are $\lambda/2$ long, have a very broad resonance band, and the region of anomalous dispersion in a dielectric utilizing such

⁵ Anomalous dispersion occurs in optical substances when the frequency of the incident light wave approximates one of the vibrational resonance frequencies of the molecule. On the long wavelength side of this resonance region the index of refraction is greater than one and increases rapidly as the resonance wavelength is approached. Dispersion, which is the change of index of refraction with frequency, is therefore very high, but it is the "normal" type of dispersion. At resonance, the absorption of the wave is high and the substance becomes opaque. At still shorter wavelengths, the resonance phenomenon acts to make the index of refraction less than its long wavelength value, often less than one, and the index again varies rapidly with wavelength, but because this is an "abnormal" type of dispersion, it is called the region of anomalous dispersion.

rods is very large, i.e., the dispersion is not very great. If it is desired to have a highly dispersive material, this band can be considerably reduced by the process of tilting the rods so that they are more nearly perpendicular to the electric vector. They thereby become "loosely coupled" to the incident wave and acquire a higher Q . Some unsymmetrical arrangement such as that shown in Fig. 14 is necessary to insure that the radiation damping of the elements is actually reduced, since a uniform tilt of all the elements would allow the array to radiate, unhindered, a wave polarized parallel to the elements. Measurements of the index of refraction of a dielectric made up of successive arrays of rods arranged as in Fig. 14 are

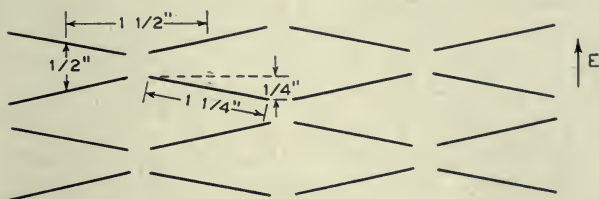


Fig. 14—Array of metal rods to produce a narrow dispersion band.

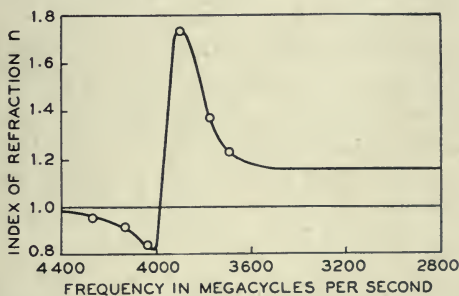


Fig. 15—Measured index of refraction of a metallic dielectric comprising the tilted rods of Fig. 14.

given in Fig. 15, and one observes a marked similarity to the behavior of the index of refraction of dielectrics in the region of their anomalous dispersion. Because of the rapid change of n with wavelength, such material may be useful as a means of separating narrow band radio channels by the use of prisms or by lenses having several feed horns located at the optimum focal points for the various frequency bands involved.

MICROWAVE REPEATER ANTENNA

For radio relay applications, there are three electrical characteristics which antennas should possess. The first is high gain (effective area), as this will reduce the path loss and accordingly the requirements on trans-

mitter power. The second is good directional qualities so as to minimize interference from outside sources and also interferences between adjacent antennas. The third is a good impedance match so that reflections between the antenna and the repeater equipment will not distort the transmitted signal. These characteristics should preferably be attainable without the imposition of severe mechanical or constructional requirements.

The shielded lens type of antenna (lens in the aperture of a horn) lends itself well to repeater work because of its moderate tolerance requirements, its good directional properties associated with the shielding, and the desirable impedance characteristic obtainable by tilting the lens. The delay lens offers the additional advantage of broad band performance with the consequent possibility of operating on several bands widely separated in wavelength. In this section, construction details and performance of a 6-foot square aperture strip type delay lens will be discussed.

(a) *Design of the Artificial Dielectric*

The operating frequency band envisioned for this antenna was 3700 to 4200 megacycles ($\lambda = 7.15$ to 8.1 cm), and to keep the element length sufficiently well removed from the half-wave resonant length a value of $\frac{3}{4}$ " for the strip width was chosen. The index of refraction of solid polystyrene is approximately 1.61 and this introduces a reflection loss (mismatch loss) at each surface of 0.225 decibels. To reduce this loss to 0.18 db. the artificial dielectric was designed to have an n of 1.5 as this still did not impart too great a thickness to the lens. A combination of strip spacings which yields an n of 1.5 was found to be $\frac{3}{8}$ " in the horizontal direction and $1\frac{5}{16}$ " center to center spacing in the vertical direction as shown in Fig. 16. The construction method of Fig. 6 was used which involved inserting .002" copper strip into slots cut in foam sheets.

(b) *Lens Design*

Several lens shapes were possible: (1) bi-convex, (2) plano-convex with the flat side toward the feed, and (3) plano-convex with the curved side toward the feed. For a given thickness and therefore weight of lens, the third possibility produces the shortest over-all structure of lens plus horn feed, and it was accordingly selected. The curved side is then a hyperboloid of revolution as shown in Fig. 17 and for the chosen focal length of 60", the profile, as calculated from the equation shown for $n = 1.5$, reaches a maximum thickness of 16". To eliminate the reflection from the lens into the feed, a lens tilt could have been employed. It was found that a quarter wave offset of one half of the lens relative to the other half could accomplish this same purpose, because reflected rays from one half of the lens then undergo a one half wavelength longer path in returning to the feed and the reflections from the two halves cancel. As this process, however, is completely effective only at one frequency, the final lens design employed both a

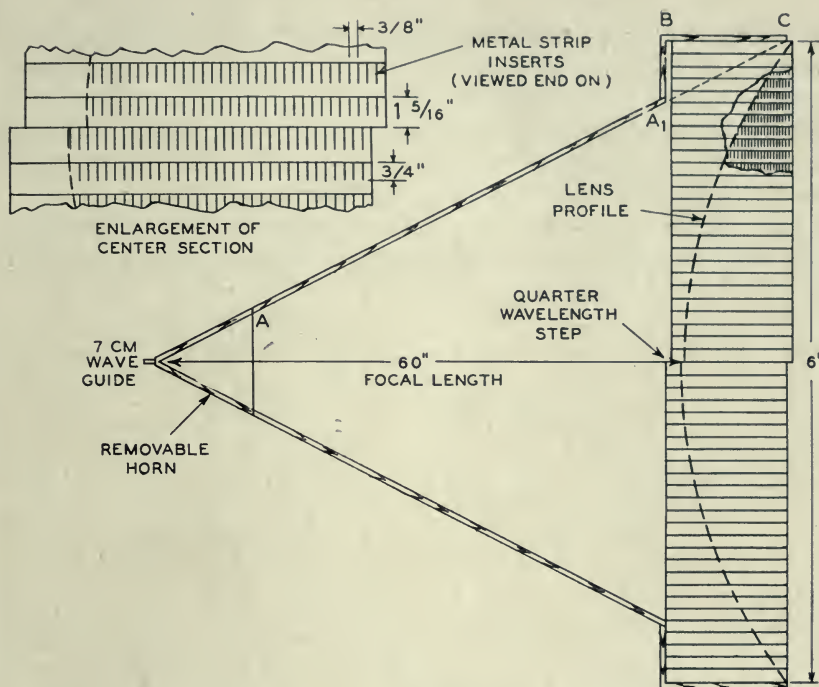


Fig. 16—Construction of a six-foot-square shielded delay lens for repeater application. The wooden horn has a metal interior surface.

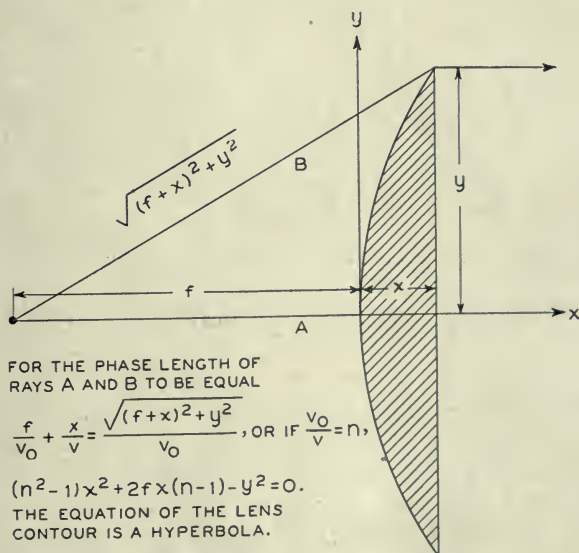


Fig. 17—Profile equation for a delay lens,

tilt in one plane and a quarter wave offset in the other plane. To utilize most efficiently the space afforded the antenna on the top of the relay tower, the lens aperture was made square. Since an unstepped lens has, by its nature, a circular aperture, the four corners of the horn aperture must be filled in with lens material as sketched in Fig. 18. The step height is designed for midband wavelength and in the present case follows the equation of Fig. 17 with the focal length reduced from the value used for the main lens section by $K\lambda/(n-1)$ where K is an integer. That integer K is

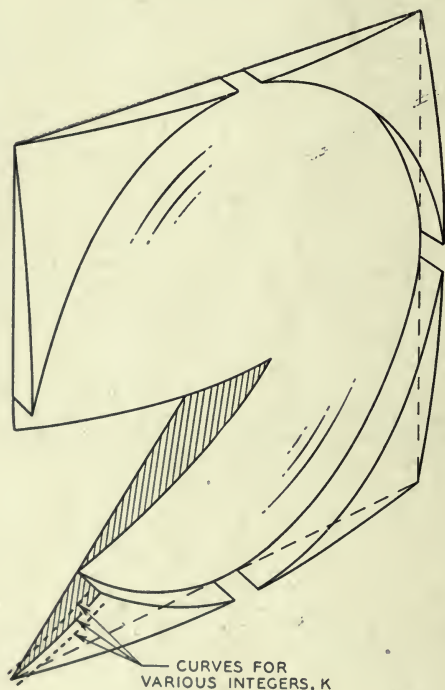


Fig. 18—Filling in the four corners of a square horn aperture with lens material.

selected which brings the step profile nearest the horn corners as indicated in Fig. 18.

(c) Horn Shield

Lenses can be energized by means of a small feed horn placed at the focal point, but to obtain the best directional properties a full metallic shield extending from the wave guide feed up to the sides of the lens should be employed. The use of foam slabs in the construction of the delay lens prevents this shield from extending completely up to the lens as indicated by the extension of $A-A_1$ (shown dotted) in Fig. 16. Optical formulas

indicate, however, that for such large apertures, the amount of diffraction (spreading of the waves outside of the pyramid formed by the dotted lines) will be small, and it is only necessary that the sides AA_1 , A_1B and BC all be conducting to insure good back lobe suppression and other desired properties associated with a horn shield.

(d) Performance

The gain of this antenna over an isotropic radiator is plotted in Fig. 19. The top curve is the theoretical gain of a uniform current sheet of the same area ($6' \times 6'$), the lower curve the gain of a $6' \times 6'$ area having 60% effective area. The points, which fall approximately on the lower curve, are the experimentally measured gain values of this antenna at the frequencies

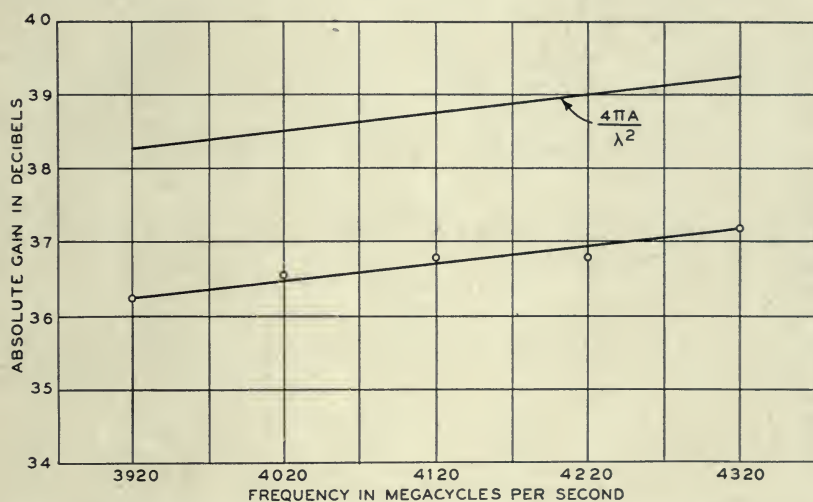


Fig. 19—Measured gain characteristics of the six-foot square shielded lens of Fig. 16. The lower line indicates 60% effective area and the circles are experimental points.

indicated. The constant percentage effective area indicates that the index of refraction of the lens material remains quite constant over the indicated frequency band. In contrast to this, the $10' \times 10'$ metal plate lenses¹ ($n < 1$) exhibit, at the band edges, a falling off of $1\frac{1}{2}$ decibels from midband gain for a 10% wavelength band.

The magnetic plane pattern of the lens when fed by a 6-inch square feed horn is shown in Fig. 20. The remarkable symmetry of the minor lobes in Fig. 20 and also in Fig. 9 shows that the phase fronts of the waves radiated by these antennas are very accurately flat. This result emphasizes the tolerance advantages of the lens over the reflector. It is believed that the

¹ Loc. cit.

further improvement in pattern of this lens over previous wave guide type metal lenses is attributable to the absence of steps, as they tend to introduce diffraction effects. The symmetry of the pattern also indicates a high degree

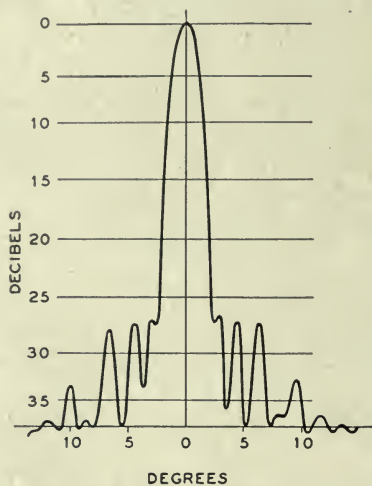


Fig. 20—Directional pattern of the lens of Figs. 8 and 16 when fed with a six-inch square electromagnetic horn.

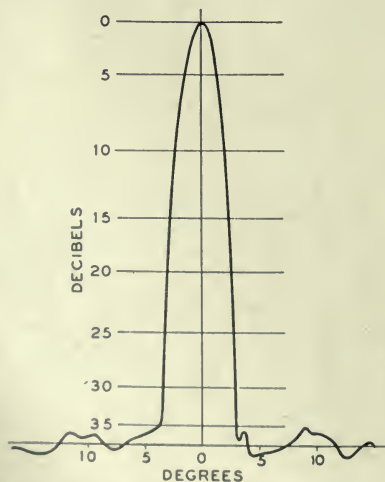


Fig. 21—Directional pattern of the same lens when enclosed with a full horn shield.

of homogeneity of the dielectric (the lens is 16 inches thick); this is a property not always shared by ordinary dielectrics such as polystyrene.

When a full horn shield is used, the illumination across the aperture has a very strong taper (somewhat stronger than a cosine taper because of the wide flare angle of the horn). This results in a very effective suppression of the close-in side lobes as shown in Fig. 21. In the vertical plane, the side

lobes are not very well suppressed because the illumination is only slightly tapered. However, for repeater work, lobes in the vertical plane are not as objectionable as lobes in the horizontal plane because interfering signals generally originate from other stations lying in the horizon plane.

The impedance match of the shielded lens antenna is affected by the discontinuity of the wave guide at the expanding horn throat and by whatever energy is reflected back into the feed line from the lens. By tuning means, the throat mismatch can be held to less than 0.2 db SWR (1.02 V.S.W.R.) over the 500 mc band and by a combination of lens tilt and quarter-wave step in the lens the SWR due to the energy reflected from the lens was also held to less than 0.2 db over the band.

PART II—THEORETICAL CONSIDERATIONS

We turn now to a consideration of the electromagnetic theory underlying the operation of artificial dielectrics. If the polarizability of the individual conducting element employed is known, equation (2) will permit a calculation of the effective dielectric constant. Before proceeding to this, however, a brief review of dipole moment, dipole potential, and dielectric polarization will be given (MKS units).

DIPOLE MOMENT, POTENTIAL AND POLARIZATION

Two charges, $+q$ and $-q$, displaced a small distance from one another, constitute an electric dipole. If the vector joining them is called $d\mathbf{s}$, the dipole moment is defined as

$$\mathbf{m} = q d\mathbf{s}. \quad (5)$$

The potential V , at any point, due to a point-charge q is defined as

$$V = q/4\pi\epsilon r, \quad (6)$$

where r is the distance from the point in question to the charge q . The potential V due to a dipole of moment $\mathbf{m}$ is

$$V = \frac{|\mathbf{m}|}{4\pi\epsilon r^2} \cos \theta, \quad (7)$$

where θ is the angle between $\mathbf{r}$ and $\mathbf{m}$.

A conducting object becomes polarized when placed in an electric field. Its dipole moment $\mathbf{m}$ depends upon the field strength $\mathbf{E}$ and upon its own polarizability α :

$$\mathbf{m} = \alpha \mathbf{E}. \quad (8)$$

If there are N elements per unit volume, the polarization $\mathbf{P}$ of the artificial dielectric is

$$\mathbf{P} = N\alpha \mathbf{E}. \quad (9)$$

But $\mathbf{P}$ is related to the displacement vector $\mathbf{D}$ and the dielectric constant as follows:

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad (10)$$

so that $\epsilon = \epsilon_0 + N\alpha$ which is equation (1). A knowledge of α thus permits a determination of ϵ . We obtain α from (8) by first finding the dipole moment $\mathbf{m}$ of the particular shape of element when immersed in a uniform field $\mathbf{E}$.

CALCULATION OF DIELECTRIC CONSTANTS OF ARTIFICIAL DIELECTRICS⁶

(1) *Conducting Sphere*

Consider a perfectly conducting sphere immersed in an originally uniform field of potential

$$V = -Ey = -Er \cos \theta. \quad (11)$$

The free charges on the sphere are displaced by the applied field and it thereby becomes a dipole whose moment $\mathbf{m}$ we wish to determine. The external potential field is the sum of the applied potential and the dipole potential, and from (7) and (11) we have

$$V_{out} = -Er \cos \theta + \frac{m \cos \theta}{4\pi\epsilon_0 r^2}. \quad (12)$$

The internal field is zero because the sphere is conducting. At a boundary between two dielectrics, there is the requirement⁷

$$V_{outside} = V_{inside}. \quad (13)$$

Equation (13) gives, at $r = a$ (the radius of the sphere),

$$-Ea \cos \theta + \frac{m \cos \theta}{4\pi\epsilon_0 a^2} = 0, \quad (14)$$

or

$$m = 4\pi\epsilon_0 Ea^3, \quad (15)$$

the dipole moment of the sphere. From (8) we see that the polarizability of the conducting sphere is accordingly $4\pi\epsilon_0 a^3$, from which equation (3) follows.

(2) *Magnetic Effects of a conducting sphere array*

The above calculations on a conducting sphere assume an electrostatic field. At microwaves, the rapidly varying fields induce eddy currents on the surface of the sphere which prevent the magnetic lines of force from penetrating the sphere. The magnetic lines are perturbed as shown in

⁶ The author is indebted to Dr. S. A. Schelkunoff for the polarizability formulas given in this memorandum.

⁷ Smythe, "Static and Dynamic Electricity", McGraw-Hill, 1939, p. 19.

Fig. 5. Now a conducting sphere in an alternating magnetic field is equivalent to an oscillating magnetic dipole,⁸ and we should observe an effective permeability for a sphere array.⁹ The magnetic dipole field is, however, opposed to the inducing field, in other words, the dipole moment is negative. Smythe⁷ shows that the magnetic polarizability of a conducting sphere of radius a in a high frequency field is

$$\alpha_m = -2\pi\mu_0 a^3. \quad (16)$$

The effective relative permeability of an array of N spheres per unit volume is therefore

$$\mu_r = 1 - 2\pi a^3 N. \quad (17)$$

The index of refraction was given in equation (4) as the square root of the dielectric constant. This is strictly true only when the permeability of the dielectric is unity. We have seen above, however, that the sphere array at microwaves possesses an effective permeability given by (17), and therefore (4) is not valid. The correct expression for n is

$$n = \sqrt{\mu_r \epsilon_r}, \quad (18)$$

and the effective refractive index of a sphere array is accordingly

$$n = \sqrt{(1 - 2\pi N a^3)(1 + 4\pi N a^3)}, \quad (19)$$

which is smaller than that given by (3) and (4). The disk and strip arrays, besides being lighter, avoid the diminishing effect on the refractive index caused by the perturbing of the magnetic lines.

(3) *Conducting Circular disk*

The determination of the dipole moment of a disk involves the use of ellipsoidal coordinates and will not be carried through here. Smythe⁷ gives an expression for the torque on a flat disk of radius a in a uniform field in Gaussian units. In M.K.S. units, his formula becomes

$$\begin{aligned} T &= 4\pi\epsilon_0 \frac{2a^3 E^2 \sin 2\theta}{3\pi} = 4\pi\epsilon_0 \frac{2a^3 E^2 2 \sin \theta \cos \theta}{3\pi} \\ &= \left[\frac{16a^3 \epsilon_0}{3} (E \sin \theta) \right] [E \cos \theta]. \end{aligned} \quad (20)$$

The first bracket represents the dipole moment, the second the field, and the product gives the torque. When the plane of the disks is parallel to E $\sin \theta$ is one, and from $m = \alpha E$, we have

$$\alpha = \frac{16\epsilon_0 a^3}{3}. \quad (21)$$

⁸ T. S. E. Thomas, *Wireless World*, Dec. 1946, p. 322.

⁹ L. Lewin, *Jour. I. E. E.*, Part III, Jan. 1947, p. 65.

⁷ Loc. Cit. Eq. 14, p. 397.

⁷ Ibid, eq. 7, p. 163.

so that, from (2),

$$\epsilon_r = 1 + \frac{16}{3} N a^3 \quad (22)$$

(4) Strips

The calculation of the dipole moment of a thin conducting strip as used in the strip lens of Fig. 15 involves two dimensional elliptic coordinates and will also be omitted here. Again, the torque is given in Smythe⁷ for an elliptic dielectric cylinder, from which we obtain,

$$\alpha = \frac{\pi \epsilon_0 s^2}{4}, \quad (23)$$

where s is the strip width, so that

$$\epsilon_r = 1 + \frac{\pi}{4} s^2 n, \quad (24)$$

where n is the number of strips per sq. unit area looking end on at the strips.

(5) Validity of the polarizability equations

Equation (2), which expresses the dielectric constant to be expected from an array of N elements each having a polarizability of α , was derived (by (8), (9) and (10)) by assuming that the field acting on an element, and tending to polarize it, was the impressed field $\mathbf{E}$ alone. This is a satisfactory assumption when the separation between the objects is so large that the elements themselves do not distort the field acting on the neighboring elements. Such is not the case when the value of ϵ_r exceeds 1.5 or thereabouts. For the usually desired values of ϵ_r of 2 or 3 it is thus seen that the above formulas such as equations (22) and (24) will yield only qualitative results and that the exact spacings of the elements to produce a desired refractive index will have to be determined by experimental methods.

For lattices having 3-dimensional symmetry, an improvement over equation (2), which takes into account not only the impressed field $\mathbf{E}$ but also the field due to the surrounding elements, is the so-called Clausius-Mosotti equation:

$$\frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} = \frac{N\alpha}{3\epsilon_0} \quad (25)$$

This, along with a similar expression for the permeability [replacing (17)], would permit a fairly accurate determination from (18) of n for the conducting sphere array.

⁷ Ibid, eq. 6, p. 97.

RESONANCE EFFECTS

It was stated earlier that the size of the elements should be small relative to a half wavelength in order for the refractive index to be independent of frequency. A qualitative idea of this criterion can be obtained by an elementary analysis of forced oscillations of dipoles. It is known that a dielectric medium which is composed of elements that resonate under the action of an alternating electric field, such as atoms having bound electrons, will exhibit a dielectric constant which varies with frequency:¹⁰

$$\epsilon_r = 1 + \frac{k}{f_0^2 - f^2}, \quad (26)$$

where f_0 is the frequency of resonance of the element, f the frequency of the incident radiation and k a proportionality constant. Thus when f is small relative to f_0 , ϵ_r is practically independent of f .

As a means of estimating the change in refractive index with frequency of metal delay lenses, we consider a specific example: Let $n = 1.50$ when the elements are $\lambda/4$ in length, i.e. when $f^2 = \frac{1}{4}f_0^2$, then the last term of (26) equals 1.25. Decreasing f by 20% reduces n from 1.50 to 1.46. Thus the change in n from midband to the edges of a $\pm 10\%$ band is about .02. From this, at 7 cm, the phase front, even for a lens 30" thick, should remain plane to within $\pm \frac{1}{10}$ wavelength over this 20% band of wavelengths. If the members had been made $\frac{1}{8}$ wavelength long, the variation in n from the design frequency all the way down to D.C. would have amounted to only 1.2%.

SUMMARY

A *metallic dielectric* is constructed by arraying conducting elements in a three-dimensional lattice structure. For electromagnetic waves whose wavelength is long compared to the size and spacing of the elements, this structure displays an effective dielectric constant and index of refraction which is sensibly constant over wide frequency bands. Lenses can be designed according to these principles which will focus microwaves and longer radio waves as a glass lens focusses light waves. Such lenses have the advantage of broad-band performance over the earlier waveguide type metal lenses and they retain the advantages of light weight over dielectric lenses. As microwave antennas, they are superior to parabolic dish reflectors from the standpoint of warping and twisting tolerance, profile tolerance, directional properties and impedance match. By eliminating the steps in the lens, the directive patterns are made cleaner and an increase in absolute

¹⁰ See for example, Joos, *Theoretical Physics*, Blackie & Son, Book 4. Chapter 4.

gain results. Because of the broad-band properties of the artificial dielectric, this improved gain can be maintained over a very wide band of frequencies. The lenses can be built to focus waves of any polarization and, if desired, the dielectric can be designed to exhibit strong dispersion. Theoretical calculations of the expected dielectric constant are in fairly good agreement with experiment for values less than 1.5; for higher values an accurate determination of the true value must be obtained experimentally.

ACKNOWLEDGEMENT

The author wishes to express his appreciation to Dr. S. A. Schelkunoff for his advice and consultation, and to the members of the Holmdel Radio Research Laboratory, especially to Mr. William Legg, for assistance and cooperation during the course of this work.

A Non-reflecting Branching Filter for Microwaves

By W. D. LEWIS and L. C. TILLOTSON

Microwave branching filters are required as integral parts of multi-channel microwave radio relay systems. These filters must have characteristics which are difficult to attain if one attempts to extend familiar lower frequency techniques to the microwave region. A novel network configuration, through which currently anticipated requirements can be met without excessive difficulty, is described in this paper.

In this configuration individual constant resistance channel dropping units are formed of appropriate assemblies of two hybrid circuits, two band reflection filters and two quarter wavelengths of line. An assembly of N channel dropping units in cascade then forms an N channel constant resistance branching network.

The mechanical and electrical characteristics of a practical five channel branching filter of this type are described. As a result of experience with this prototype filter it can be stated with some safety that these requirements can be fulfilled with a network of this type. Experimentally observed impedance, insertion loss and phase characteristics were fully satisfactory. In addition the circuit appears to be flexible enough both electrically and mechanically to fulfill the various types of systems needs which may be encountered at branch points or when channels must be added or interchanged.

INTRODUCTION

PRESENT plans for point-to-point communication by means of microwave radio relay systems call for the operation of several radio channels between each pair of repeaters. A proposed frequency plan for the 4030 mc common carrier band (3700 to 4200 mc) specifies channels 20 mc wide spaced 40 mc center to center. A possible arrangement for a five-channel radio repeater station is illustrated in Fig. 1. This arrangement is calculated to utilize the available frequency space in an efficient and technically sound manner.

If this channel disposition is to be achieved without a costly increase in the number of antennas and the size of the supporting towers, radio frequency branching networks must be provided which connect the individual transmitting or receiving circuits at each repeater point to a common antenna (Fig. 1). If this connection is to be made without excessive loss of power these branching devices must have adequate adjacent channel rejection, low ohmic loss, and good impedance match in the channel bands. An excellent impedance match is especially desirable if circuit disturbances resulting from echoes in the long waveguide lines which lead from the filter assemblies to the antennas are to be minimized (Fig. 1).

Since the type of microwave radio repeater now planned¹ obtains most of its gain at intermediate frequencies, the IF amplifiers will reject all spurious

¹ See H. T. Friis, "Microwave Repeater Research", to appear in the April 1948 issue of *B. S. T. J.*

signals entering the receiver except those in the vicinity of the receiver image bands. Because of this, suppression requirements on the branching filters, except possibly in the vicinity of receiver image bands, are not severe.

Problems connected with the design of suitable microwave branching filters naturally differ considerably from previous filter problems. The discrimination and band utilization requirements are readily met, but the

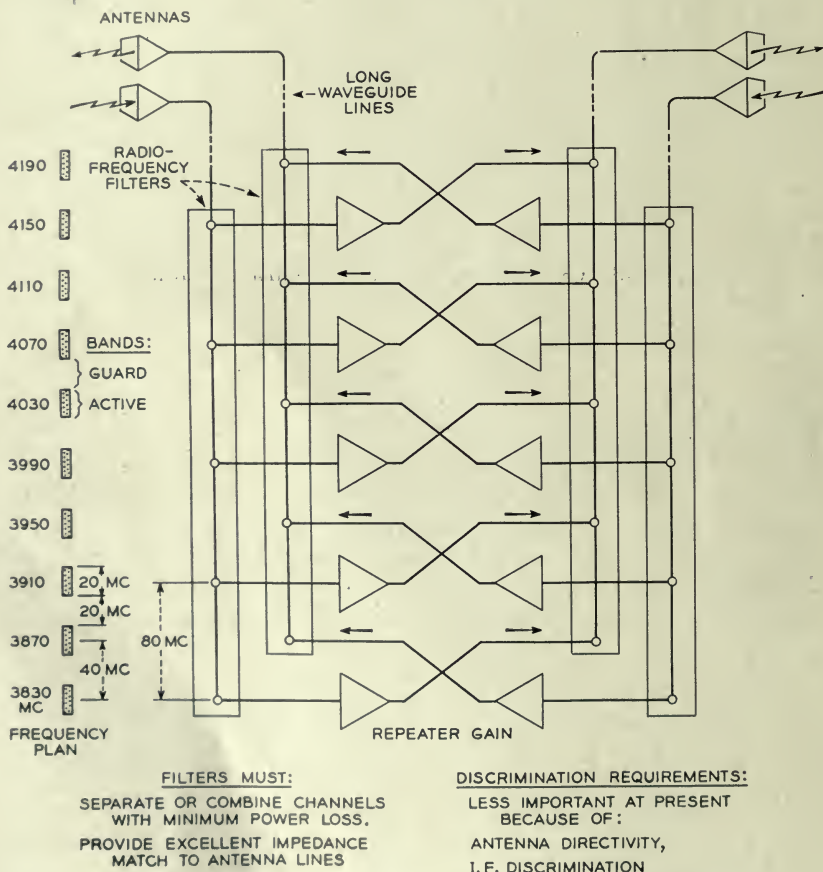


Fig. 1—Possible five-channel radio repeater station schematic diagram.

impedance control required is not easy to obtain by familiar filter techniques. Not more than about 5% variation in impedance or about 0.5 db standing wave ratio in the channels could be tolerated. At lower frequencies channel passing networks which can be connected in series or parallel to form a channel branching filter can be designed on the basis of lumped circuit theory and built of coils, condensers and resistances, but in the microwave region

simple parallel or series connections and simple lumped circuit elements do not exist. Although in the interests of flexibility it would be desirable to add or substitute individual channels without affecting other channels in a branching filter, the convenient possibility of doing so at a high impedance level on vacuum tube grids is not yet available in the microwave region.

A satisfactory two-channel waveguide branching filter has been constructed following partially 'classical' methods. This filter, designed for the New York-Boston experimental radio relay system, is composed of two channel-passing cavity filters each connected to a common input line through one arm of an E plane Y junction, the waveguide analogue of a series connection (Fig. 2). This solution, although relatively simple where only two channels are required, becomes quite complex when more than two are involved, since in every channel the sum of the interactions of all the inactive filters on

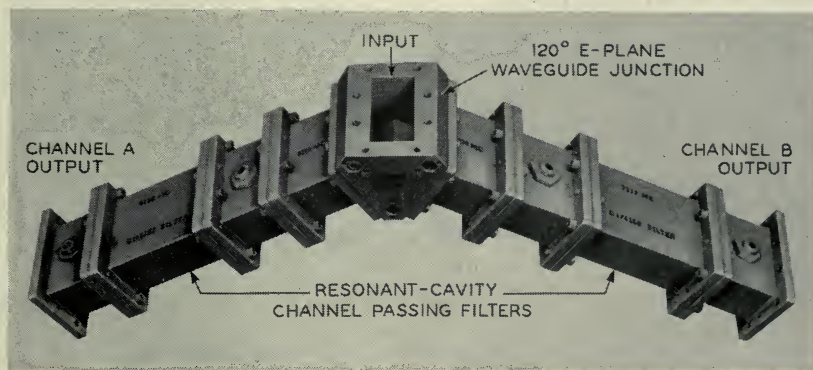


Fig. 2—Branching filter for New York-Boston experimental radio relay system.

transmission through the active filter must be zero. It is evidently not easy to satisfy this condition, particularly since in doing so one must take accurate account of the change with frequency of the effective phase shift of all waveguide connecting lines. And even if such a solution were found it would be valid for only one set of channels, so that the problem must be solved all over again for every change in channel arrangement.

As a result of these difficulties and after a few attempts to overcome them, it became apparent that a more flexible method of microwave filter construction should be found. Constant resistance filters, which provide discrimination by absorbing or diverting the unwanted incident waves rather than by reflecting them, can be useful in any frequency range. In the microwave region, where the shortest connecting pipe may be many wavelengths long, constant resistance devices are particularly helpful. Accordingly a constant resistance channel-dropping network was devised which

could extract one channel from the line, while permitting others to pass through it without disturbance. Several of these networks were then placed in cascade to make up the required filter.

THE HYBRID CHANNEL-DROPPING UNIT

After several possibilities were considered the constant resistance channel-dropping circuit illustrated schematically in Fig. 3 was selected.² This circuit is made up of two hybrid junctions, two identical channel reflection filters tuned to the dropped channel, and two quarter wavelength sections of line.

In order to understand the operation of this circuit, the properties of a hybrid circuit, Fig. 4(a), must first be understood. This circuit, which has

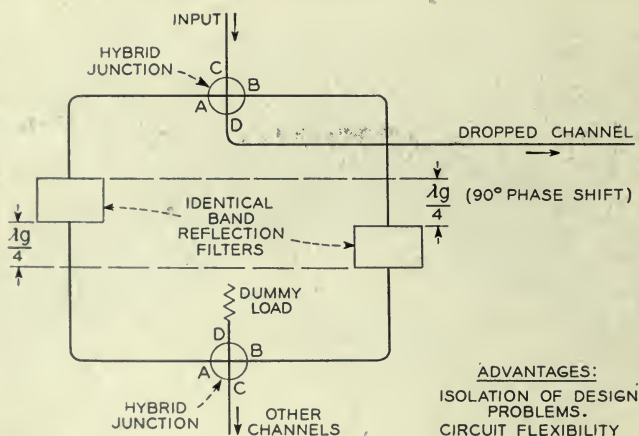


Fig. 3—Possible constant impedance channel dropping filter.

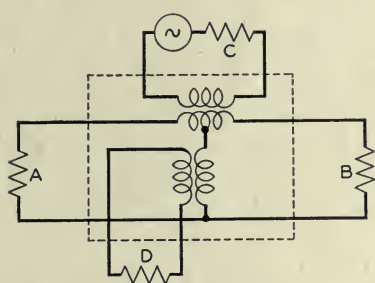
been embodied at voice frequencies as the hybrid coil and at microwaves as the hybrid junction (E-H plane T junction)³ can be represented schematically as in Fig. 4(b). Let us connect four transmission lines, A, B, C, D, each terminated in its characteristic impedance to the four pairs of terminals of the hybrid circuit. Then if each of these lines is matched to the pair of terminals to which it is connected, the following characteristics will result. Power in transmission line C flowing towards the junction will divide equally into lines A and B, flow away from the junction and be absorbed in loads A and B. None of this power will appear in line D or be reflected back into

² Lumped element networks with properties similar to those of this circuit have been devised by Vos and Laurent, U. S. Patent 1,920,041, and Bobis, U. S. Patent 2,044,047. A. G. Fox of these laboratories has independently devised similar microwave circuits.

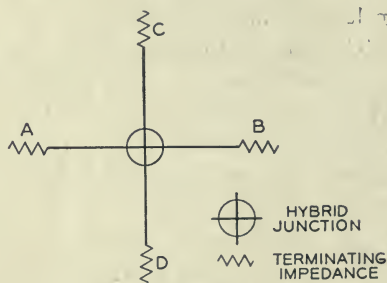
³ W. A. Tyrrell, Hybrid Circuits for Microwaves, Proc. I.R.E., Vol. 35, pp. 1294-1306, November 1947.

line C. Similarly, power in line D flowing towards the junction, will appear equally in A and B but not in C or back in D. If these characteristics hold, then by the principle of reciprocity similar characteristics must hold for lines A and B. If proper planes of reference are chosen this behavior can be described in a slightly different, but equivalent, manner. If waves in both A and B flow towards the junction the vector sum of the voltages of these times a constant (0.707) appears in C and the vector difference times the same constant appears in D, but nothing is reflected back into A or B. An equivalent statement can be made if the waves start in C and D.

With the properties of the hybrid in mind, and if it is assumed that the hybrids, the identical reflection filters and the quarter wave lines are perfect and free of ohmic loss, the operation of the circuit of Fig. 3 is easy to understand, and is as follows: A wave entering from arm C of the input hybrid is divided equally into the two arms A and B. None of the power in this wave



(a) CLASSICAL HYBRID COIL



(b) HYBRID WAVEGUIDE JUNCTION SCHEMATIC DIAGRAM

Fig. 4—*a.* Classical hybrid coil.
b. Hybrid junction schematic diagram.

is reflected back into the arm C or appears initially in arm D. The two equal components of the wave now travel along the lines which are connected to the arms A and B of the input hybrid. If the frequency lies outside the band of the reflection filters the waves travel through these filters and appear in phase in the arms A and B of the output hybrid. The vector sum of these two waves appears in the arm C of the output hybrid and has an amplitude equal to that of the original input wave. Consequently all energy in the input line incident on this network, except that lying in the band of the reflection filters, will pass through it to the output line.

If now the frequency of the input wave lies within the band of the reflection filters, the two equal components of the wave traveling away from the input hybrid will be reflected at the filters and will travel back towards the input hybrid. One of these components must, however, travel twice through an extra quarter wavelength of line, and will therefore be reversed in phase; with

respect to the other component when it reappears at the input hybrid. The two components will consequently combine in the fourth or difference arm of the input hybrid to form a wave equal to the original input wave.

The circuit of Fig. 3 is only one of a general class of hybrid filter circuits. From one viewpoint these circuits resemble spectroscopes, and, from another, ordinary lumped circuits. We could, for example, replace the band reflection circuits of Fig. 3 with any two identical four-terminal networks. This circuit would still retain its constant resistance character. One of its branches would contain all energy reflected by the four terminal networks; the other would contain all energy passed by them.

In particular if the two identical filters are of the channel passing type, input waves within this channel will be transmitted by both filters and will combine at the output hybrid and appear in arm C. All of the other channels will be reflected by the filters and thus will appear in arm D of the input hybrid, provided that the assumed 90° phase shift holds over a band which includes all of the channels.

The particular configuration of Fig. 3 was chosen to minimize the effect of practical limitations. The dropped channel width (20 mc) is only a small fraction of its midband frequency (about 4000 mc). Consequently when band reflection filters are used in Fig. 3, the change with frequency of the nominally quarter wave sections of guide does not seriously affect the performance of the filter and the impedance match of arm D of the hybrids needs to be good only in the vicinity of the dropped channel.

The circuit shown in Fig. 3 is a constant resistance network which drops the channel corresponding to the reflection filter. Several in sequence as indicated in Fig. 5 constitute a constant resistance channel branching filter. In the sections to follow we will give an account of the physical configuration and electrical performance of a branching network designed according to this pattern to meet the radio frequency requirements of a typical practical radio relay system containing five 20 mc radio frequency channels, spaced 80 mc center to center.

THE WAVEGUIDE HYBRID

In choosing a hybrid configuration which could be successfully used in the network of Fig. 5, both electrical and mechanical requirements were considered. Since frequency f_n passes through $2n - 1$ hybrids it is evident that if acceptable overall performance is to be obtained, the balance and impedance characteristics of each hybrid must be excellent. A broad-band balance can be obtained with relative ease by attaching the 'driven' arms (A and B on Fig. 5) symmetrically. Fortunately the strict impedance requirement applies to only one of the two 'driving' arms (C and D), the other being required to transmit only a single channel.

Because of the reentrant nature of the circuit of Fig. 5 it is evidently desirable, if not essential, to employ a hybrid with a convenient mechanical layout. If, for example, a familiar E-H plane junction type of hybrid were used in combination with waveguide filters built in a straight piece of waveguide, twenty-four E or H plane right angle waveguide bends would be required in the construction of each six-channel branching network. To avoid this extra complication and expense a hybrid configuration with 'driven' output arms parallel to the well matched 'driving' input arm was sought.

The hybrid configuration settled upon was the one shown in Fig. 6. Here the electrical and geometrical requirements discussed above are met simultaneously by connecting the driving arm C to the symmetrically located driven arms A and B through a smoothly tapered E plane Y junction. The taper is approximately one half wavelength long in the center of the 3700–4200

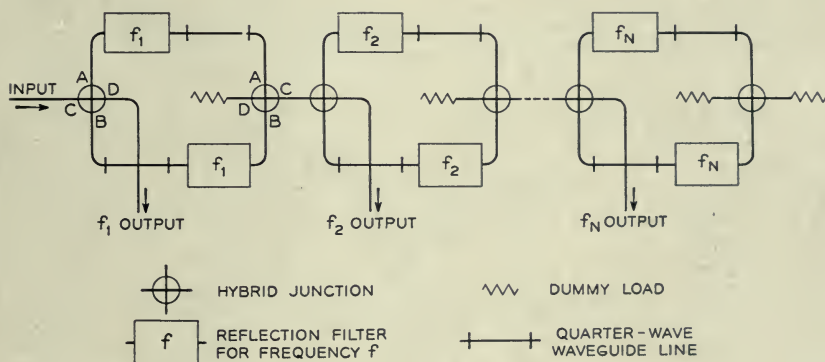


Fig. 5—N Channel branching filter.

mc band. The driving arm D is connected to arms A and B through a coaxial line. This line is coupled to waveguide D by means of a conventional probe. The center conductor traverses the Y junction space in such a way that it is normal to the electric vectors of guides A, B, and C, and is effectively coupled to A and B but not to C by means of a probe P fastened through it normally (Fig. 6).

THE BAND REFLECTION FILTERS

The ideal reflection filter for the circuit of Figs. 3 and 5 would reflect perfectly within a certain band and pass perfectly outside of this band. However, the ratio of bandwidth to band spacing in a given branching filter (20 mc to 80 mc) is such that a sufficiently good approximation to this ideal can be obtained in theory if each reflection filter employs only three resonances, Fig. 7(a). These could be effectively series resonant circuits placed at quarter

wave intervals along the guide, properly distributed in impedance level and all tuned close to the center frequency of the channel to be extracted, Fig 7(b). The practical question was to find how to obtain these resonances in an easily constructed and adequately adjustable form.

Early experiments indicated that a probe inserted in the broad side of the guide far enough so that its end formed an appreciable capacitance with the opposite side could be made to resonate in a series resonant fashion. Impedance levels available through this means of coupling were, however, far lower than required. Accordingly an alternate method in which the probe is

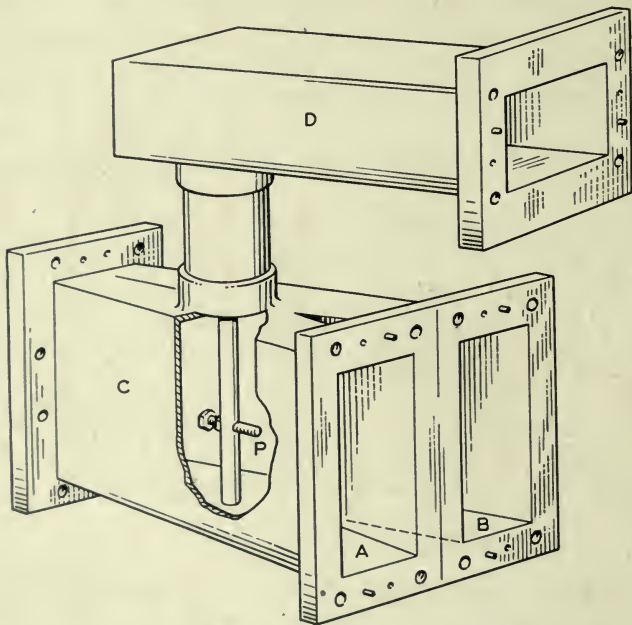


Fig. 6—Hybrid junction.

inserted in the narrow side of the guide was selected (Fig. 8). The probe is made long enough to approach the opposite narrow wall of the guide and is tipped by a capacitive disk. The capacity of the disk and consequently the resonant frequency of the circuit is adjusted by means of a screw in the wall just opposite the disk. Since this rod is inserted perpendicular to the narrow guide wall it is normally uncoupled to the principal mode in the guide. An adjustable coupling is achieved by inserting a screw in the broad side of the guide just above the probe. Insertion of this screw disturbs the symmetry of the field and couples the rod to the guide. Increasing the screw insertion increases the coupling and consequently varies the impedance level of the

equivalent series resonant circuit continuously from infinity down to any value within the range required.

It should be pointed out that impedance and frequency adjustments are not in practice completely independent, but do at least permit the realization of any required value. Furthermore, mutual coupling between the probes interferes with the exact realization of the circuits of Fig. 7. This coupling does not interfere, however, with the realization of satisfactory filter characteristics.

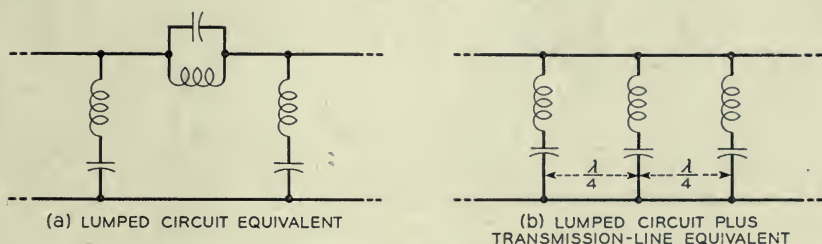


Fig. 7—Band reflection filter.

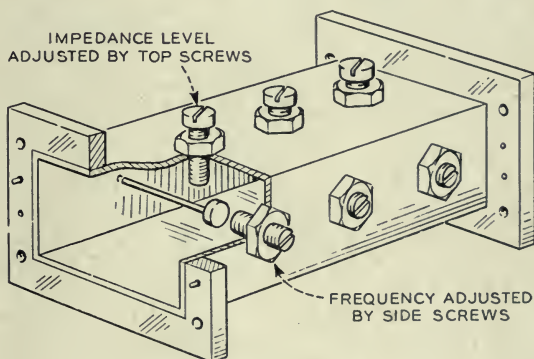


Fig. 8—Waveguide band reflection filter.

When the reflection filter of Fig. 7 was embodied in waveguide form through use of the series resonant circuits just described, the configuration illustrated in Fig. 8 resulted. It was found that, with the exception of a slight change in disk to guide wall spacing, a single configuration could be tuned to any of the desired channels.

ELECTRICAL PERFORMANCE

Individual components as described above were constructed, adjusted and assembled to form a five-channel branching network (Fig. 9). The electrical performance of this network was trimmed systematically, then was measured

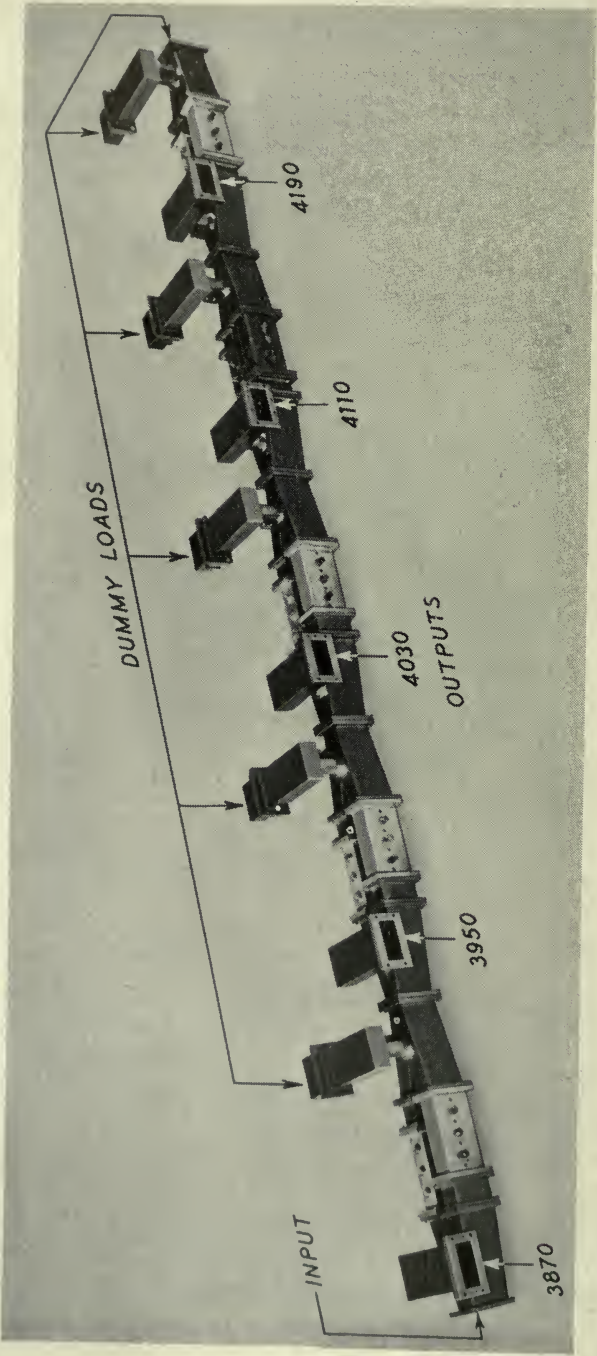


Fig. 9—Five-channel branching filter.

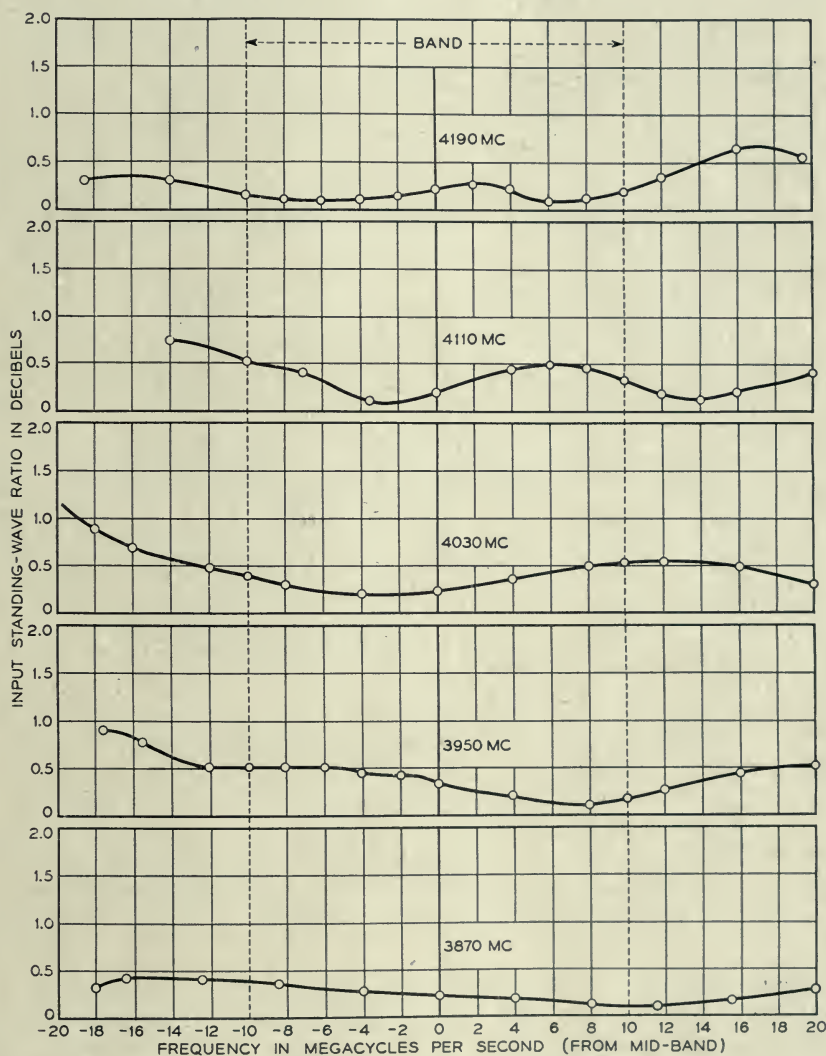
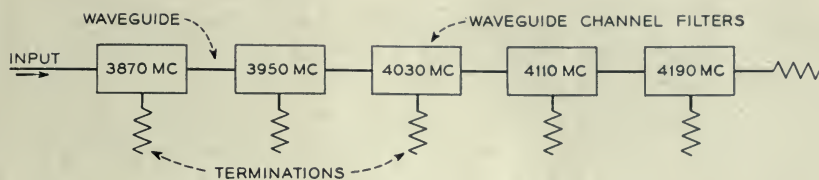


Fig. 10—Input standing wave ratio of hybrid branching filter.

point by point with a double detection measuring set. With all outputs terminated the standing wave ratio observed at the input line was under 0.6 db in

all channels. This quantity is plotted in Fig. 10. The insertion loss measured between the input line and the various output lines varied from about 0.5 db in the lowest frequency channel closest to the input to about 1.0 db in the highest frequency channel farthest from the input. This loss is plotted in Fig. 11. Since 0.5 db was approximately the loss observed in each individual

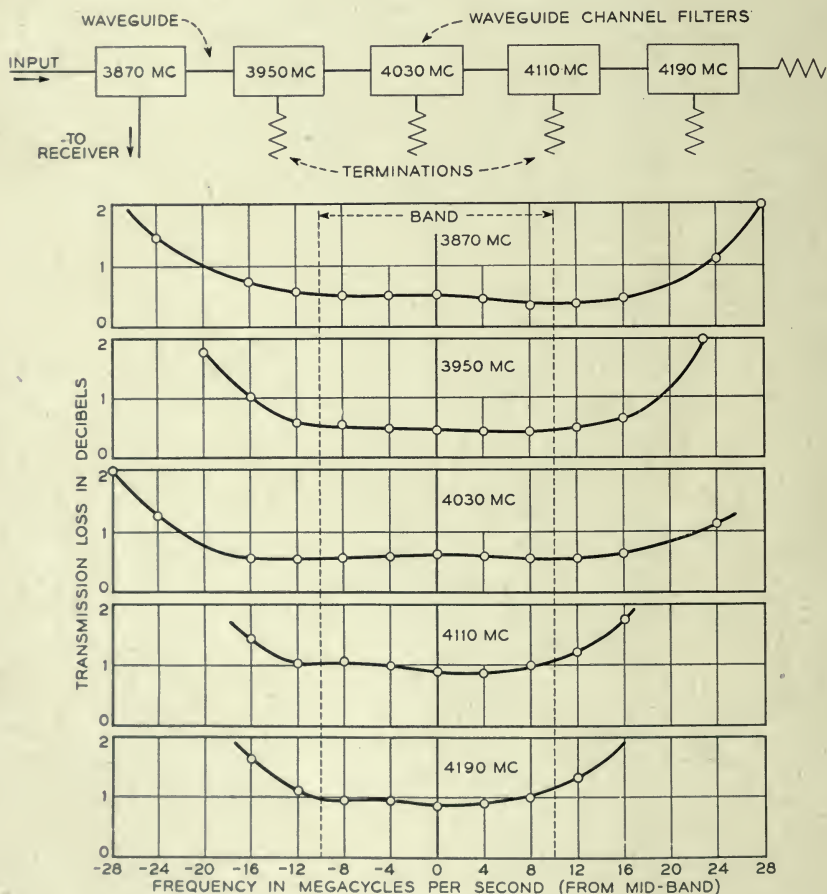


Fig. 11—Transmission loss of hybrid branching filter.

channel dropping unit, it appears logical to assume that the progressive increase in loss is due to passage of the higher frequencies through the lower frequency circuit components.

The measured discrimination against other channels and image responses was 20 db or more. This modest amount of selectivity is sufficient since the primary function of the present filter is branching, and hence the amount of

discrimination needed is only that required to prevent a significant energy loss in the channel bands. If crosstalk considerations indicate that more r-f discrimination is required, this can be obtained by placing auxiliary filters in the branch arms. This can be done without complication since the branching filter is a constant resistance device.

The delay distortion was examined but was found to be less than the errors of measurement, i.e. 2 millimicroseconds.

ACKNOWLEDGEMENTS

The authors, in carrying out the work described in this paper, have been helped by discussions with many people. Particular credit is due to J. R. Pierce for a variety of stimulating suggestions, and to their colleagues at the Holmdel Radio Laboratory for much fundamental work on microwave transmission.

A Note on a Parallel-Tuned Transformer Design

By V. C. RIDEOUT

An analysis of the parallel-tuned transformer used in radio-frequency amplifiers has been made for a slightly over-coupled case. The resulting design formulas are simple and practical.

Two cases are discussed: (a) the so-called matched transformer, with resistance loading on each side; (b) a transformer with loading on one side only, which has the same pass-band and phase characteristics as the matched transformer, but gives 3 db more gain when used as an interstage.

A special arrangement of (a) where the matched transformer design is used with one resistor removed, giving a transformer with a considerably double-humped pass-band characteristic and about 6 db more gain, is also discussed.

INTRODUCTION

THE parallel-tuned transformer has been used in radio-frequency amplifiers for many years.¹ In an excellent paper,² Christopher based design formulas on the principles of the broad band filter. In other publications, simple circuit analyses have been used and design formulas have been based on the assumption of a small ratio of bandwidth to mid-frequency which often fails to be adequate when the wide bands required for modern television and multiplexing services are encountered.

A transitionally flat transformer design may be obtained by setting the first three derivatives of the absolute value of the transfer impedance, with respect to frequency, equal to zero. The resulting design formulas are somewhat unwieldy.

The transformer design to be described here is based upon two simple circuit conditions³ applied to the fundamental case of a parallel-tuned transformer with resistance loading on each side:

I. Both sides of the transformer are tuned to the same frequency.

II. The transmission loss is zero at the tune frequency. (This condition is responsible for the term "matched transformer" used to describe this case).

The resulting transformer has a slightly double-humped characteristic with less than 0.005 db dip for a coupling coefficient of 0.5. Because of the slight overcoupling this transformer design gives a little more gain and bandwidth than the critically coupled case. Its main advantage lies in the fact that simpler design formulas can be used.

¹ H. T. Friis and A. G. Jensen, *High Frequency Amplifiers*, *B. S. T. J.*, Vol. III, April 1924, pp. 181-205.

² A. J. Christopher, *Transformer Coupling Circuits for High-Frequency Amplifiers*, *B. S. T. J.*, Vol. XI, Oct. 1932, pp. 608-621.

³ This analysis is based on an old unpublished report by H. T. Friis.

DERIVATION OF DESIGN FORMULAS

(a) Matched, or Symmetrically Loaded Transformer.

In the circuit of Fig. 1:

$$\left. \begin{aligned} E_1 &= I_1(R_1 + 1/j\omega C_1) + I_2(-1/j\omega C_1) \\ 0 &= I_1(-1/j\omega C_1) + I_2(j\omega L_1 + 1/j\omega C_1) + I_3(-j\omega M) \\ 0 &= I_2(-j\omega M) + I_3(j\omega L_2 + 1/j\omega C_2) + I_4(-1/j\omega C_2) \\ 0 &= I_3(-1/j\omega C_2) + I_4(R_2 + 1/j\omega C_2) \end{aligned} \right\} \quad (1)$$

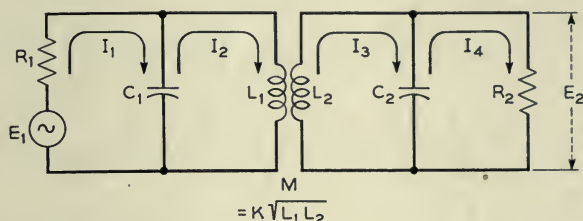


Fig. 1.—Parallel tuned, matched, or symmetrically loaded transformer.

Eliminating I_1 , I_2 and I_3 among these equations gives:

$$\begin{aligned} E_1/I_4 &= \frac{L_1 R_2 + L_2 R_1}{M} + \omega^2 \left[\left(M - \frac{L_1 L_2}{M} \right) (C_1 R_1 + C_2 R_2) \right] \\ &+ j \left(\omega^3 \left[\left(M - \frac{L_1 L_2}{M} \right) C_1 C_2 R_1 R_2 \right] + \omega \left[\frac{R_1 R_2}{M} (L_1 C_1 + L_2 C_2) \right. \right. \\ &\quad \left. \left. - \left(M - \frac{L_1 L_2}{M} \right) \right] - \frac{1}{\omega} \cdot \frac{R_1 R_2}{M} \right) \end{aligned} \quad (2)$$

If the coefficient of mutual coupling is k , and the first and second inductance and capacitance are resonant at ω_1 and ω_2 radians per second, respectively, we can substitute in (2) the expressions,

$$M = k\sqrt{L_1 L_2}, L_1 = 1/\omega_1^2 C_1, L_2 = 1/\omega_2^2 C_2, I_4 = E_2/R_2 \quad (3)$$

This gives the general expression,

$$\begin{aligned} E_1/E_2 &= \frac{R_1 \sqrt{C_1 C_2}}{k} \left\{ \left[\frac{\omega_1}{\omega_2} \cdot \frac{1}{R_2 C_2} + \frac{\omega_2}{\omega_1} \cdot \frac{1}{R_1 C_1} \right] \right. \\ &\quad \left. - \omega^2 \left[\frac{1 - k^2}{\omega_1 \omega_2} \left(\frac{1}{R_2 C_2} + \frac{1}{R_1 C_1} \right) \right] \right. \\ &\quad \left. + j \left(\frac{-\omega^3}{\omega_1 \omega_2} (1 - k^2) + \omega \left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + \frac{1 - k^2}{\omega_1 C_1 R_1 \omega_2 C_2 R_2} \right] \right. \right. \\ &\quad \left. \left. - \frac{\omega_1 \omega_2}{\omega} \right) \right\} \end{aligned} \quad (4)$$

The application of circuit condition I, ($\omega_1 = \omega_2 = \omega_0$) to (4) gives,

$$E_1/E_2 = \frac{R_1 \omega_0 \sqrt{C_1 C_2}}{k} \left\{ \left(\frac{1}{\omega_0 C_1 R_1} + \frac{1}{\omega_0 C_2 R_2} \right) - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \left(\frac{1}{\omega_0 C_1 R_1} + \frac{1}{\omega_0 C_2 R_2} \right) - j \left[\left(\frac{\omega}{\omega_0} \right)^3 (1 - k^2) - \frac{\omega}{\omega_0} \left(2 + \frac{1 - k^2}{\omega_0^2 C_1 C_2 R_1 R_2} \right) + \frac{\omega_0}{\omega} \right] \right\} \quad (5)$$

Circuit condition II (zero transmission loss at $\omega = \omega_0$) is satisfied if $|E_1/E_2| = 2 \sqrt{R_1/R_2}$. Substituting this condition in (5) and solving for k gives:

$$k^2 = \frac{\omega_0^2 C_1 C_2 R_1 R_2 + 1 \pm j(\omega_0 C_1 R_1 - \omega_0 C_2 R_2)}{(\omega_0 C_1 R_1)^2 + (\omega_0 C_2 R_2)^2 + (\omega_0^2 C_1 C_2 R_1 R_2)^2 + 1} \quad (6)$$

Since k must be real,

$$\omega_0 C_1 R_1 = \omega_0 C_2 R_2 \quad (7)$$

and from (6) and (7)

$$\omega_0 C_1 R_1 = \omega_0 C_2 R_2 = \sqrt{1 - k^2}/k \quad (8)$$

From (5) and (8) the transmission formula for this transformer is,

$$\frac{E_1}{E_2} = \frac{\sqrt{1 - k^2}}{k^2} \sqrt{\frac{C_2}{C_1}} \left[\frac{2k}{\sqrt{1 - k^2}} \left(1 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \right) - j \left(\left(\frac{\omega}{\omega_0} \right)^3 (1 - k^2) - \frac{\omega}{\omega_0} (2 + k^2) + \frac{\omega_0}{\omega} \right) \right] \quad (9)$$

or, $\frac{E_1}{E_2} = \left| \frac{E_1}{E_2} \right| e^{j\phi}$, where,

$$\left| \frac{E_1}{E_2} \right| = \frac{\sqrt{1 - k^2}}{k^2} \left(\frac{C_2}{C_1} \right)^{\frac{1}{2}} \left\{ \frac{4k^2}{1 - k^2} \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \right]^2 + \left[\left(\frac{\omega}{\omega_0} \right)^3 (1 - k^2) - \frac{\omega}{\omega_0} (2 + k^2) + \frac{\omega_0}{\omega} \right]^2 \right\}^{\frac{1}{2}} \quad (10)$$

$$\phi = \arctan \frac{\sqrt{1 - k^2}}{2k} \cdot \frac{(\omega/\omega_0)^4 (1 - k^2) - (\omega/\omega_0)^2 (2 + k^2) + 1}{(\omega/\omega_0)^3 (1 - k^2) - \omega/\omega_0} \quad (11)$$

The transmission loss, $\mathcal{L}$, defined as the ratio of the maximum available power from the generator to the output power may be obtained from (10),

$$\mathcal{L} = \frac{R_2}{4R_1} \left| \frac{E_1}{E_2} \right|^2 = \left\{ \frac{1}{k^2} \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \right]^2 + \frac{1 - k^2}{4k^4} \left[\left(\frac{\omega}{\omega_0} \right)^3 (1 - k^2) - \frac{\omega}{\omega_0} (2 + k^2) + \frac{\omega_0}{\omega} \right]^2 \right\} \quad (12)$$

or the loss in decibels equals $10 \log \mathcal{L}$.

The transmission delay, δ , may be obtained from (11).

$$\delta = \frac{d\phi}{d\omega} = \frac{\left(\frac{\omega}{\omega_0} \right)^4 (1 - k^2)^2 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2)^2 + 4k^2 - 1 + \left(\frac{\omega_0}{\omega} \right)^2}{\mathcal{L}} \quad (13)$$

$$\cdot \frac{\sqrt{1 - k^2}}{2k^3\omega_0}$$

seconds.

The input impedance seen from the generator terminals is,

$$Z_{in} = R_1 \left\{ \frac{\left(\frac{\omega}{\omega_0} \right) k \sqrt{1 - k^2} - j \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \right]}{\left[\frac{\sqrt{1 - k^2} \left(\frac{\omega_0}{\omega} \right)}{k} \right] \left[-1 + 2 \left(\frac{\omega}{\omega_0} \right)^2 - \left(\frac{\omega}{\omega_0} \right)^4 (1 - k^2) \right] - j \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 (1 - k^2) \right]} \right\} \quad (14)$$

The formulas (12), (13) and (14) give the important characteristics of this transformer. Table I gives expressions for these characteristics in the general case, and for $k = 0.5$, for several picked frequencies. Of these frequencies ω_0 and $\omega_0/\sqrt{1 - k^2}$ are the frequencies of zero transmission loss, $\omega_0/\sqrt{1 - k^2}$ is the geometric midband, and $\omega_0/\sqrt{1 + k}$ and $\omega_0/\sqrt{1 - k}$ are the cut-off frequencies of an infinite filter made of identical transformers of this type.

The input impedance, transmission loss, and the transmission delay are plotted against ω/ω_0 for a matched transformer with $k = 0.5$ in Fig. 2. From these curves, and from Table I the following important characteristics of this transformer will be noted:

(1) The pass-band is approximately symmetrical with frequency, rather than with the logarithm of frequency, as in circuits which have a low-pass analog.

TABLE I

Relative Frequency ω/ω_0	Transmission Loss (db)		Delay $\times \omega_0$ (radians)		Input Impedance	Input Standing Wave Ratio in db, $k = 0.5$
	General Case	$k = 0.5$	General Case	$k = 0.5$		
1	0	0	$\frac{2\sqrt{1-k^2}}{k}$	3.464	R_1	0
$1/\sqrt{1-k^2}$	0	0	$\frac{2\sqrt{1-k^2}}{k}$	3.464	R_1	0
$1/\sqrt[4]{1-k^2}$	$\frac{0.136k^4}{2-k^2}$ approx.	0.0048	$\frac{3 + \sqrt{1-k^2}}{4 + k^4/8(2-k^2)} \cdot \frac{2\sqrt{1-k^2}}{k}$ approx.	3.344	$R_1 \left(\frac{k\sqrt[4]{1-k^2} - j}{1 - \sqrt{1-k^2}} \cdot \frac{1}{2 - \sqrt{1-k^2}} - j \right)$	0.585
$1/\sqrt{1+k}$	$10 \log \frac{5-k}{4}$	0.52	$\frac{6-k}{5-k} \cdot \frac{2\sqrt{1-k^2}}{k}$	4.234	$R_1(1 + j\sqrt{1-k})$	6.02
$1/\sqrt{1-k}$	$10 \log \frac{5+k}{4}$	1.39	$\frac{6+k}{5+k} \cdot \frac{2\sqrt{1-k^2}}{k}$	4.094	$R_1(1 - j\sqrt{1+k})$	10.06

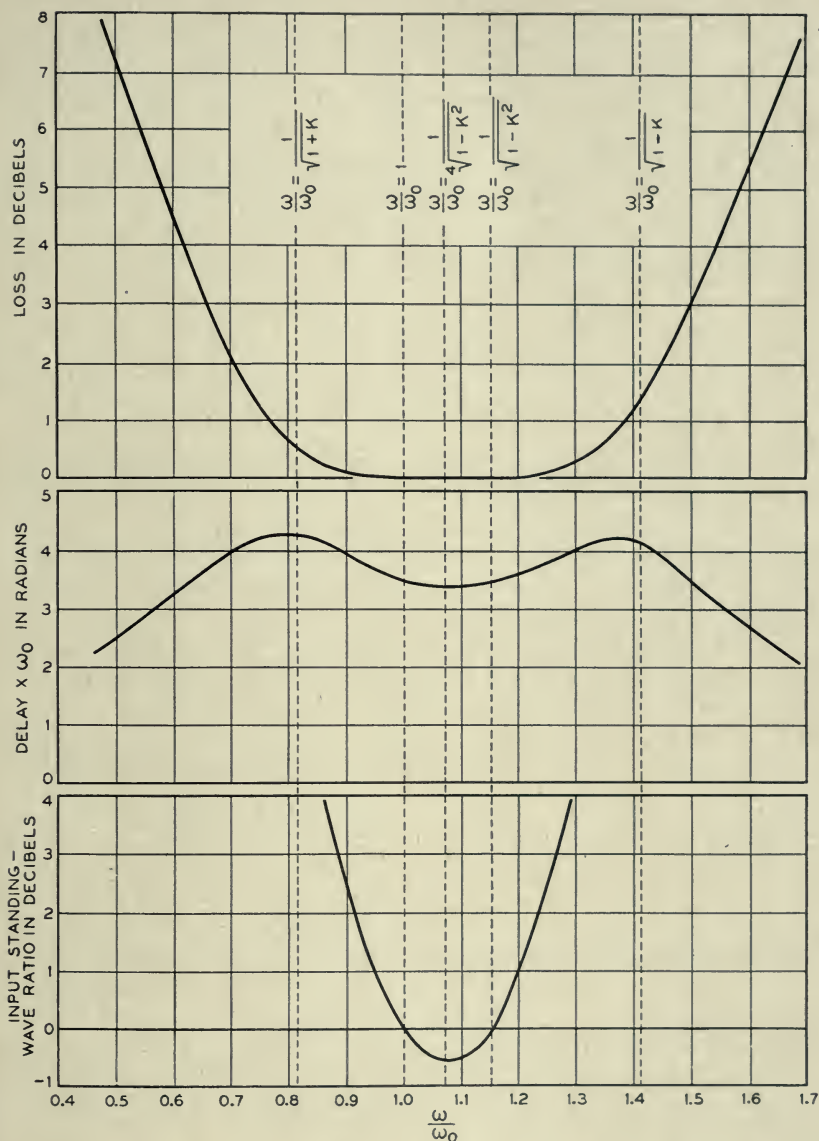


Fig. 2.—Loss, delay, and input standing-wave ratio versus relative frequency for the matched transformer for $k = 0.5$. Loss and delay curves also apply to the case of the mis-matched transformer for $k = 0.632$.

(2) The delay curve is also quite symmetrical. Circuits which can be derived from low-pass forms have more delay on the low-frequency side, where loss increases faster. The symmetry of the delay curve makes phase compensation easier.

(3) The input standing-wave ratio passes through zero on either side of mid-band, and although it has a 0.6 db dip at mid-band for the case of $k = 0.5$, it is under one db over a larger frequency range than in the transitionally coupled case.

(4) The bandwidth between points one db down on the loss curve is very nearly equal to the bandwidth between the cut-off frequencies. This bandwidth is simply related to the tune frequency f_0 and the geometric midband frequency f_m as shown below.

$$\left. \begin{aligned} \Delta f &= f_0/\sqrt{1-k} - f_0/\sqrt{1+k} \\ &\cong f_0 k/\sqrt{1-k^2} \\ &\cong f_m k/\sqrt{1-k^2} \end{aligned} \right\} \quad (15)$$

Thus, given required values of f_m , Δf , C_1 (or R_1), and C_2 (or R_2), the matched transformer can be designed by the use of formulas (3), (8) and (15). These formulas are summarized at the end of this paper.

A transformer of different phase and loss characteristics results if the loading resistance is removed from one side of a matched transformer. Equation (5), if $R_2 = \infty$, becomes,

$$\frac{E_1}{E_2''} = \frac{\sqrt{1-k^2}}{k^2} \sqrt{\frac{C_2}{C_1}} \left[\frac{k}{\sqrt{1-k^2}} \left(1 - \left(\frac{\omega}{\omega_0} \right)^2 (1-k^2) \right) - j \left(\left(\frac{\omega}{\omega_0} \right)^3 (1-k^2) - 2 \frac{\omega}{\omega_0} + \frac{\omega_0}{\omega} \right) \right] \quad (16)$$

At $\omega = \omega_0$,

$$|E_1/E_2''|^2 = C_2/C_1 \quad (17)$$

while from (10) the corresponding ratio for the matched transformer is,

$$|E_1/E_2|^2 = 4C_2/C_1 \quad (18)$$

Thus this mis-matched transformer has 6 db more gain at the tune frequency than the corresponding matched transformer, whether used as an interstage, an output or an input transformer.

The transmission loss referred to the loss at $\omega = \omega_0$ is,

$$\begin{aligned} \mathcal{L}'' &= \frac{C_1}{C_2} \left| \frac{E_1}{E_2''} \right|^2 = \frac{1}{k^2} \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 (1-k^2) \right]^2 \\ &\quad + \frac{1-k^2}{k^4} \left[\left(\frac{\omega}{\omega_0} \right)^3 (1-k^2) - 2 \frac{\omega}{\omega_0} + \frac{\omega_0}{\omega} \right]^2 \end{aligned} \quad (19)$$

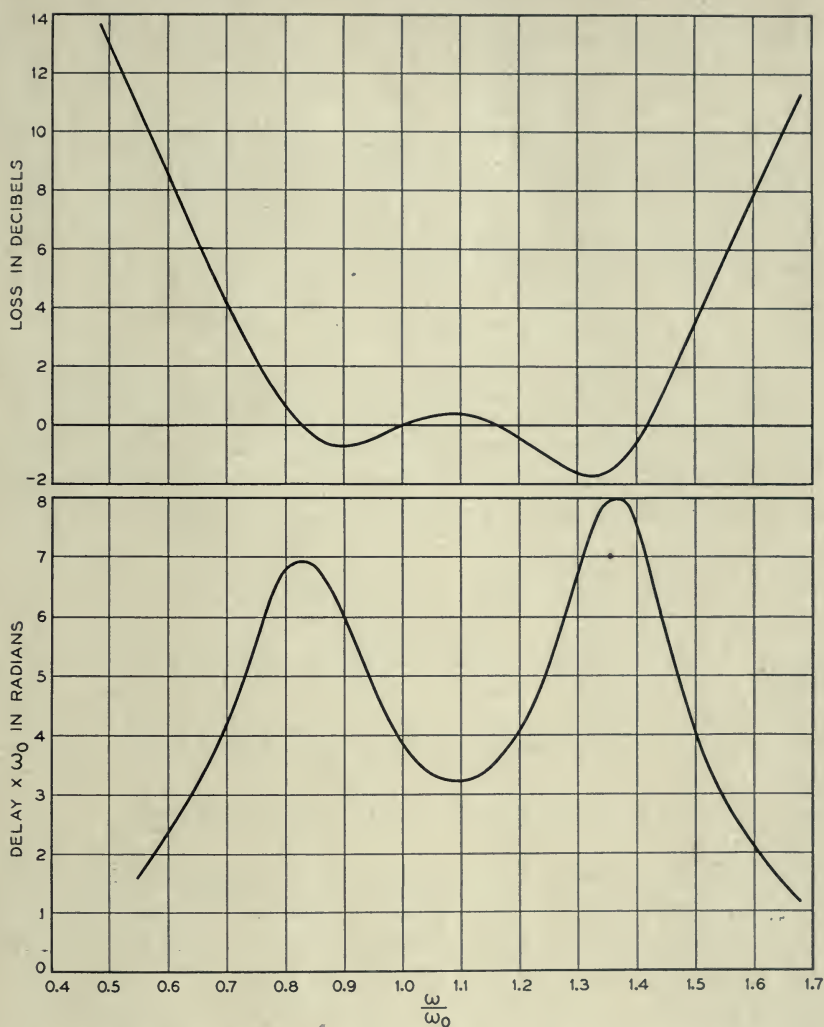


Fig. 3.—Loss and delay versus relative frequency for the mis-matched transformer obtained by making R_2 in Fig. 1 infinite, for $k = 0.5$.

The transmission delay δ' in this case is,

$$\delta' = \frac{\sqrt{1 - k^2}}{\omega_0 k^3} \frac{\left(\frac{\omega}{\omega_0}\right)^4 (1 - k^2)^2 - \left(\frac{\omega}{\omega_0}\right)^2 (1 - k^2) + (3k^2 - 1) + \left(\frac{\omega_0}{\omega}\right)^2}{\mathcal{Q}''} \text{ seconds.} \quad (20)$$

The loss is equal at four points,

$$\omega = \omega_0, \quad \omega = \omega_0/\sqrt{1-k^2}, \quad \omega = \omega_0/\sqrt{1-k}, \quad \omega = \omega_0/\sqrt{1+k}.$$

The curves of Fig. 3 show the loss and delay for this mis-matched transformer for the case where $k = 0.5$. The double-humped band-pass characteristic of the mis-matched transformer can be compensated for in an amplifier by de-tuning slightly to equalize the humps and using a single-tuned circuit in another interstage in which,

$$Q = \omega_m CR \cong \sqrt{1/k^2 - 3/4 - 3k^2/32} \quad (21)$$

The above combination of a mis-matched transformer and a single-tuned transformer, in successive interstages, will give approximately 6 db more gain than if two matched transformers were used. The actual gain obtained will depend upon the ratio of the input and output capacitances. This gain advantage carries with it the penalty of increased sensitivity to tube capac-

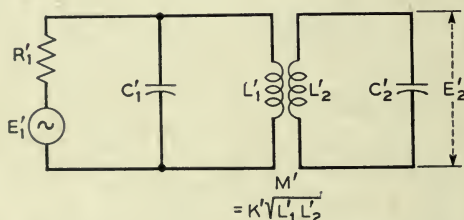


Fig. 4.—Mismatched or unsymmetrically loaded transformer.

itance variations. However, the mis-matched transformer may be used to advantage in the first interstage of an amplifier to minimize the effect of the second tube on the signal-to-noise ratio. It may also be used in the last interstage to offset any lack of power-handling capacity in the second-tube.

(b) *Mismatched or Unsymmetrically Loaded Transformer*

A transformer with only one loading resistor, as in Fig. 4, can always be designed to have the same bandpass and phase characteristics as the matched transformer⁴ shown in Fig. 1. No power transfer will occur, but the ratio of output voltage to the square root of the available input power can be determined, and set equal to this ratio for the matched transformer, except

⁴ This important principle has been described by S. Darlington in a paper: Synthesis of Reactance 4-Poles, *Journal of Mathematics and Physics*, Vol. XVIII, Sept. 1939, pp. 257-353. Independently, this principle was demonstrated to the author by W. J. Albersheim in an unpublished memorandum dealing with the transitionally flat transformer.

for a constant factor. For the matched case this ratio is, from equation (9),

$$\sqrt{\frac{E_1^2/4R_1}{E_2^2}} = \left| \frac{1}{k} \sqrt{\frac{C_2}{C_1 R_1}} - \left(\frac{\omega}{\omega_0}\right)^2 \frac{1-k^2}{k} \sqrt{\frac{C_2}{C_1 R_1}} \right. \\ \left. - j \left(\frac{\omega}{\omega_0}\right)^3 \frac{(1-k^2)\sqrt{1-k^2}}{2k^2} \sqrt{\frac{C_2}{C_1 R_1}} \right. \\ \left. + j \left(\frac{\omega}{\omega_0}\right) \frac{(2+k^2)\sqrt{1-k^2}}{2k^2} \sqrt{\frac{C_2}{C_1 R_1}} \right. \\ \left. - j \left(\frac{\omega}{\omega_0}\right) \frac{\sqrt{1-k^2}}{2k^2} \sqrt{\frac{C_2}{C_1 R_1}} \right| \quad (22)$$

For the mismatched transformer the corresponding ratio can be obtained from equation (4) by letting R_2 approach infinity, giving,

$$\sqrt{\frac{E_1'^2/4R_1'}{E_2'^2}} = \left| \frac{1}{2k'} \frac{\omega_2'}{\omega_1} \sqrt{\frac{C_2'}{C_1' R_1'}} - \frac{\omega^2}{\omega_1' \omega_2'} \right. \\ \left. \cdot \frac{1-k'^2}{2k'} \sqrt{\frac{C_2'}{C_1' R_1'}} - j \frac{\omega^3}{\omega_1' \omega_2'} \frac{1-k'^2}{2k'} \sqrt{R_1' C_1' C_2'} \right. \\ \left. + j \omega \left(\frac{\omega_2'}{\omega_1} + \frac{\omega_1'}{\omega_2'} \right) \frac{\sqrt{R_1' C_1' C_2'}}{2k'} - j \frac{\omega_1' \omega_2'}{\omega} \frac{\sqrt{R_1' C_1' C_2'}}{2k'} \right| \quad (23)$$

The two transformers are to have the same phase characteristics, and the same band-pass characteristics except for a constant factor N where,

$$N = \sqrt{\frac{E_1^2/4R_1}{E_2^2}} \div \sqrt{\frac{E_1'^2/4R_1'}{E_2'^2}} \quad (24)$$

The terms involving voltages in (22) and (23) can be eliminated by combining with (24). Equating the coefficients of like powers of ω gives (25), below.

$$\left. \begin{aligned} \frac{1}{k} \sqrt{\frac{C_2}{C_1 R_1}} &= \frac{1}{2k'} \sqrt{\frac{C_2'}{R_1' C_1'}} N \frac{\omega_2'}{\omega_1'} \\ \frac{1-k^2}{k} \sqrt{\frac{C_2}{C_1 R_1}} \frac{1}{\omega_0^2} &= \frac{1-k'^2}{2k'} \sqrt{\frac{C_2'}{C_1' R_1'}} \left(\frac{N}{\omega_1' \omega_2'} \right) \\ \frac{1-k^2}{k} \sqrt{R_1 C_1 C_2} \cdot \frac{1}{\omega_0^2} &= \frac{1-k'^2}{k'} \sqrt{R_1' C_1' C_2'} \left(\frac{N}{\omega_1' \omega_2'} \right) \\ \frac{2+k^2}{k} \sqrt{R_1 C_1 C_2} &= \frac{\sqrt{R_1' C_1' C_2'}}{k'} \left(\frac{\omega_2'}{\omega_1'} + \frac{\omega_1'}{\omega_2'} \right) N \\ \frac{\sqrt{R_1 C_1 C_2}}{k} \omega_0^2 &= \frac{\sqrt{R_1' C_1' C_2'}}{k'} (\omega_1' \omega_2' N) \end{aligned} \right\} \quad (25)$$

The last three equations were simplified by the use of equation (8). Equations (25) yield the five relations below:

$$\left. \begin{aligned} R_1' C_1' &= R_1 C_1 / 2 \\ \omega_1 &= \omega_0 \\ \omega_2 &= \omega_0 / \sqrt{1 + k^2} \\ k' &= \frac{\sqrt{2} k}{\sqrt{1 + k^2}} \\ N &= 2 \sqrt{\frac{C_2}{C_2'}} \end{aligned} \right\} \quad (26)$$

SUMMARY OF DESIGN FORMULAS

A transformer must usually meet certain requirements as to bandwidth and mid-frequency. Loss curves for various values of k , plotted against relative frequency may be used. To a very close approximation the geometric mean frequency $f_m = \omega_m / 2\pi$, may be used for the mid-band frequency, and Δf , the bandwidth between cut-off points, may be used for the bandwidth between one db points. Then,

$$\omega_m = \omega_0 / \sqrt[4]{1 - k^2} \quad (27)$$

And, from (15),

$$k \cong \frac{\Delta f / f_0}{\sqrt{1 + (\Delta f / f_0)^2}} \quad (28)$$

The other necessary relations for the matched transformer and for the mismatched transformer with the same transmission characteristics, as derived from equations (3), (8) and (26) are given below.

(a) Matched Transformer.

$$\left. \begin{aligned} \frac{1}{\sqrt{L_1 C_1}} &= \omega_0 \\ \frac{1}{\sqrt{L_2 C_2}} &= \omega_0 \\ \frac{M}{\sqrt{L_1 L_2}} &= k \\ R_1 C_1 = R_2 C_2 &= \frac{\sqrt{1 - k^2}}{\omega_0 k} \end{aligned} \right\} \quad (29)$$

(b) Mismatched Transformer:

$$\left. \begin{aligned} \frac{1}{\sqrt{L'_1 C'_1}} &= \omega_0 \\ \frac{1}{\sqrt{L'_2 C'_2}} &= \frac{\omega_0}{\sqrt{1+k^2}} \\ \frac{M'}{\sqrt{L'_1 L'_2}} &= k' = \frac{\sqrt{2}k}{\sqrt{1+k^2}} \\ R'_1 C'_1 &= \frac{\sqrt{1-k^2}}{2\omega_0 k} \end{aligned} \right\} \quad (30)$$

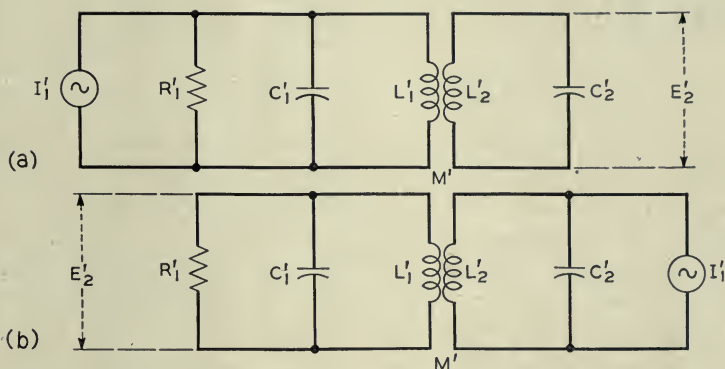


Fig. 5.—Mismatched transformer (a) fed by constant current generator, (b) with input and output reversed.

The ratio of the output voltages in the two cases is given by:

$$\left| \frac{E'_2}{E_2} \right|^2 = N^2 \left(\frac{E_1'^2 / 4R'_1}{E_1'^2 / 4R'_1} \right) = 4 \frac{R_1 C_2}{R'_1 C'_2} \left| \frac{E'_1}{E_1} \right|^2 \quad (31)$$

The relative gains of the two transformers, as given by equation (31), will be examined for three important cases:

(1) *Input Transformer*—In this case $R_1 = R'_1$, and $C_2 = C'_2$, so that from (31), the mismatched transformer gives four times, or 6 db more gain.

(2) *Output Transformer*—The voltage source in Fig. 4 may be replaced by the constant current source $I'_1 = E'_1 / R'_1$, by Norton's Theorem, as shown in Fig. 5(a). By the Reciprocity Theorem we may exchange the positions of the current source and the output voltage as shown in Fig. 5(b), without changing the value of the latter. This may also be done for the

matched transformer. The resultant circuit is that for a pentode output transformer in each case, where the plate resistance of the tube is so high compared to circuit impedance that we can consider it to be a constant current generator of $g_m E_g$ amperes, where g_m is the grid-plate transconductance and E_g the grid signal voltage. Thus in this case the mis-matched transformer also gives 6 db more gain than the matched transformer.

(3) *Interstage Transformer*—If the voltage generators of Figs. 1 and 4 are replaced by their equivalent current generators $g_m E_g = E_1/R_1$ and $g'_m E'_g = E'_1/R'_1$, then (31) becomes,

$$\left| \frac{E'_2}{E_2} \right|^2 = 4 \frac{R'_1 C_2}{R_1 C'_2} \left| \frac{g'_m E'_g}{g_m E_g} \right|^2 \quad (32)$$

The first relation in (26) may be substituted in (32) giving,

$$\left| \frac{E'_2}{E_2} \right|^2 = 2 \left| \frac{C_1 C_2 g'_m E'_g}{C'_1 C'_2 g_m E_g} \right|^2 \quad (33)$$

Thus, in the interstage case, this mismatched transformer has only 3 db more gain than the matched transformer.

Here, as in the case of the first mismatched transformer discussed in this paper, there is an increased sensitivity to capacity variations. In practice it appears that this is not always serious in wide-band intermediate-frequency amplifiers. Some improvement in noise figure and power handling may be obtained, as before, without encountering difficulties with double-peaked pass-band characteristics.

The design formulas given in this paper have been used successfully in the design of experimental intermediate frequency amplifiers in the 65 mc region having band widths of the order of 10 to 20 mc.

Statistical Properties of a Sine Wave Plus Random Noise

By S. O. RICE

INTRODUCTION

IN SOME technical problems we are concerned with a current which consists of a sinusoidal component plus a random noise component. A number of statistical properties of such a current are given here. The present paper may be regarded as an extension of Section 3.10 of an earlier paper,¹ "Mathematical Analysis of Random Noise", where some of the simpler properties of a sine wave plus random noise are discussed.

The current in which we are interested may be written as

$$\begin{aligned} I &= Q \cos qt + I_N \\ &= R \cos (qt + \theta) \end{aligned} \tag{3.4}$$

where Q and q are constants, t is time, and I_N is a random (in the sense of Section 2.8 of Reference A) noise current. When the second expression involving the envelope R and the phase angle θ is used, the power spectrum of I_N is assumed to be confined to a relatively narrow band in the neighborhood of the sine wave frequency $f_q = q/(2\pi)$. This makes R and θ relatively slowly (usually) varying functions of time.

In Section 1, the probability density and cumulative distribution of I are discussed. In Section 2, the upward "crossings" of I (i.e., the expected number of times, per second, I increases through a given value I_1), are examined.

The probability density and the cumulative distribution of R are given in Section 3.10 of Reference A. The crossings of R are examined in Section 4 of the present paper.

The statistical properties of θ' , the time derivative of the phase angle θ , are of interest because the instantaneous frequency of I may be defined to be $f_q + \theta'/(2\pi)$. The probability density of θ' is investigated in Section 5 and the crossings of θ' in Section 6. θ' is a function of time which behaves somewhat like a noise current and may accordingly be considered to consist of an infinite number of sinusoidal components. The problem of determining the "power spectrum" $W(f)$ of θ' , i.e., the distribution of the mean square value of the components as a function of frequency, is attacked in

¹ *B.S.T.J.* 23 (1944), 282-332 and 24 (1945), 46-156. This paper will be called "Reference A".

Sections 7 and 8. The correlation function of θ' is expressed in terms of exponential integrals in Section 7, the power spectrum of I_N being assumed symmetrical and centered on f_q . In Section 8, values of $W(f)$ are obtained for the special case in which the power spectrum of I_N is centered on f_q and is of the normal-law type.

It is believed that some of the material presented here may find a use in the study of the effect of noise in frequency modulation systems. For example, the curves in Section 8 yield information regarding the noise power spectrum in the output of a primitive type of system. Also, the procedure employed to obtain the expression (5.7) for $\bar{\theta}'$ may be used to show that if

$$Q\cos[(A/\omega_0)\cos\omega_0t + qt] + I_N = R\cos(qt + \theta)$$

the sinusoidal component of $d\theta/dt$ is²

$$-A(1 - e^{-\rho})\sin\omega_0t$$

where ρ is the ratio $Q^2/(2\bar{I}_N^2)$. This illustrates the "crowding effect" of the noise. The statistical analysis associated with R and θ of equations (3.4) (when the sine wave is absent) is similar to that used in the examination of the current returned to the sending end of a transmission line by reflections from many small irregularities distributed along the line. This suggests another application of the results.

ACKNOWLEDGMENT

I am indebted to a number of my associates for helpful discussions on the questions studied here. In particular, I wish to thank Mr. H. E. Curtis for his suggestions regarding this subject. As in Reference A, all of the computations for the curves and tables have been performed by Miss M. Darville. This work has been quite heavy and I gratefully acknowledge her contribution to this paper.

1. PROBABILITY DISTRIBUTION OF A SINE WAVE PLUS RANDOM NOISE

A current consisting of a sine wave plus random noise may be represented as

$$I = Q\cos qt + I_N \quad (1.1)$$

where Q and q are constants, t is the time, and I_N is a random noise current. The frequency, in cycles per second, of the sine wave is $f_q = q/(2\pi)$. In all

² The first person to obtain this result was, I believe, W. R. Young who gave it in an unpublished memorandum written early in 1945. He took the output of a frequency modulation limiter and discriminator to be proportional to either the signal frequency or to the instantaneous frequency (assumed to be distributed uniformly over the input band) of I_N according to whether Q is greater or less than the envelope of I_N . His memorandum also contains results which agree well with several obtained in this paper.

our work we denote the power spectrum of I_N by $w(f)$ and its correlation function by $\psi(\tau)$. The mean square value of I_N is denoted by ψ_0 .

The study of the probability distribution of I is essentially a study of the integral³

$$p(I) = \frac{1}{\pi\sqrt{\psi_0}} \int_0^\pi \varphi \left[\frac{I - Q \cos \theta}{\sqrt{\psi_0}} \right] d\theta \quad (1.2)$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (1.3)$$

and $p(I)$ is the probability density of I , i.e. $p(I)dI$ is the probability that a value of current selected at random will lie in the interval $I, I + dI$. Another expression for $p(I)$ is given by equation (3.10-6) of Reference A, namely

$$p(I) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-izI - \psi_0 z^2/2} J_0(Qz) dz \quad (1.4)$$

where $J_0(Qz)$ denotes the Bessel function of order zero.

The substitutions

$$y = \frac{I}{\sqrt{\psi_0}}, \quad a = \frac{Q}{\sqrt{\psi_0}} \quad (1.5)$$

enable us to write (1.2) as

$$p_1(y) = \sqrt{\psi_0} p(I) = \frac{1}{\pi} \int_0^\pi \varphi(y - a \cos \theta) d\theta, \quad (1.6)$$

where $p_1(y)$ denotes the probability density of y . This is the expression actually studied. Curves showing $p_1(y)$ and the cumulative distribution function

$$\begin{aligned} \int_{-\infty}^I p(I_1) dI_1 &= \int_{-\infty}^y p_1(y_1) dy_1 \\ &= \frac{1}{\pi} \int_0^\pi \varphi_{-1}(y - a \cos \theta) d\theta, \end{aligned} \quad (1.7)$$

where

$$\varphi_{-1}(x) = \int_{-\infty}^x \varphi(x_1) dx_1 = \frac{1}{2} + \frac{1}{2} \operatorname{erf} (x/\sqrt{2}) \quad (1.8)$$

³ W. R. Bennett, "Response of a Linear Rectifier to Signal and Noise," *Jour. Acous. Soc. Amer.* Vol. 15 (1944), 164-172, and *B.S.T.J.* Vol. 23 (1944), 97-113.

are shown in Figs. 1 and 2. The curves for $a = 10$ and $a = \sqrt{10}$ were computed by Simpson's rule from (1.6) and (1.7), and the curves for $a = 1$ were computed from the series (1.10) given below. Since both $\varphi(x)$ and

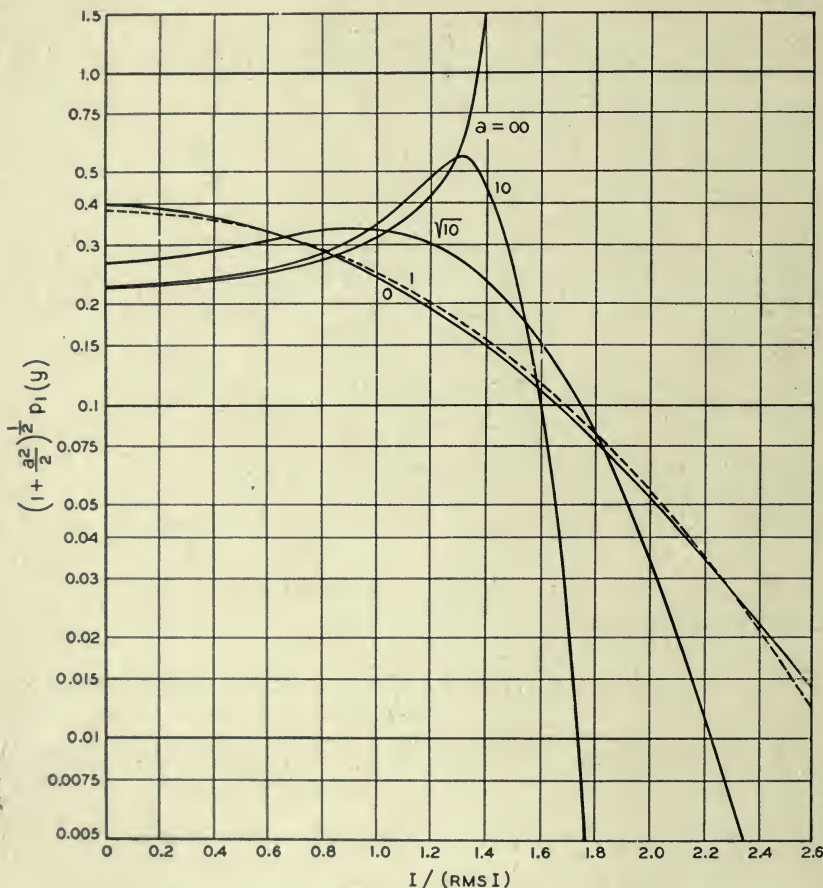


Fig. 1—Probability density of sine wave plus noise.

$I = Q \cos qt + I_N$, $a = Q/\sqrt{\psi_0}$, $y = I/\sqrt{\psi_0}$, $\sqrt{\psi_0} = \text{rms } I_N$

$p_1(y) dI/\sqrt{\psi_0}$ = probability total current lies between I and $I + dI$
 $y(1 + a^2/2)^{-1/2} = I/(\text{rms } I)$. Curves are symmetrical about $y = 0$.

$\varphi_{-1}(x)$ are tabulated⁴ functions the integrals in (1.6) and (1.7) are well suited to numerical evaluation.

⁴ $\varphi(x)$ is given directly and $\varphi_{-1}(x)$ may be readily obtained from W.P.A., "Tables of Probability Functions," Vol. II, New York (1942). The functions $\varphi^{(m)}(y)$ are tabulated in Table V of "Probability and its Engineering Uses" by T. C. Fry (D. Van Nostrand Co., 1928).

The form assumed by $p_1(y)$ as the parameter a becomes large is examined in the latter portion (from equation (1.12) onwards) of the section.

Series which converge for all values of a but which are especially suited for calculation when $a \leq 1$ may be obtained by inserting the Taylor's series (in powers of x) for $\varphi(y+x)$ and $\varphi_{-1}(y+x)$, $x = -a \cos \theta$, in (1.6) and (1.7) and integrating termwise. When we introduce the notation⁴

$$\varphi^{(m)}(y) = \frac{d^m}{dy^m} \varphi(y) = \frac{1}{\sqrt{2\pi}} \frac{d^m}{dy^m} e^{-y^2/2} \quad (1.9)$$

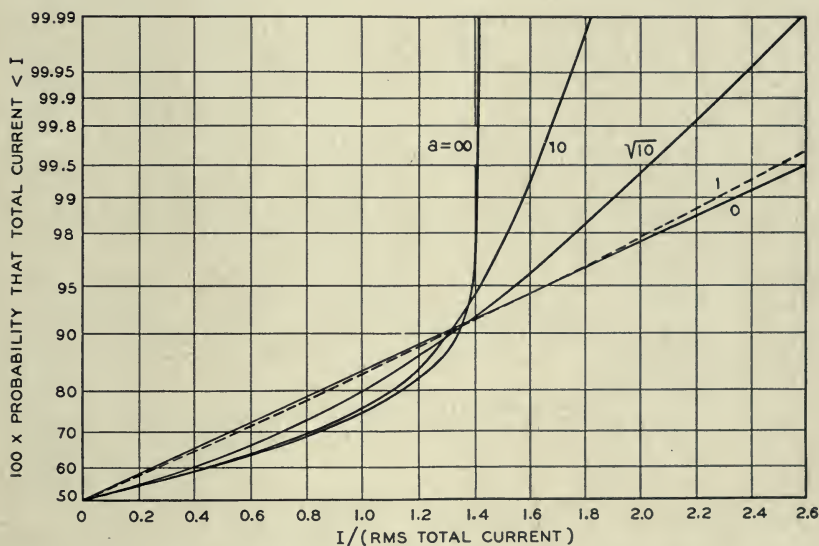


Fig. 2—Cumulative distribution of sine wave plus noise.

Ordinate = $100 \int_{-\infty}^y p_1(y_1) dy_1$. See Fig. 1 for notation.

we obtain

$$p_1(y) = \sum_{n=0}^{\infty} \frac{1}{n!n!} \left(\frac{a}{2}\right)^{2n} \varphi^{(2n)}(y) \quad (1.10)$$

$$\int_{-\infty}^y p_1(y_1) dy_1 = \varphi_{-1}(y) + \sum_{n=1}^{\infty} \frac{1}{n!n!} \left(\frac{a}{2}\right)^{2n} \varphi^{(2n-1)}(y)$$

The second equation of (1.10) may be shown to be valid by breaking the interval $(-\infty, y)$ into $(-\infty, 0)$ and $(0, y)$. In the first part,

$$\int_{-\infty}^0 p_1(y_1) dy_1 = \varphi_{-1}(0)$$

since both sides have the value $1/2$. In the second, term by term integration is valid since the series integrated are uniformly convergent as may be seen from the inequality

$$|\varphi^{(n)}(y)| \leq \left(\frac{n!}{2\pi}\right)^{1/2} \left(\frac{2}{\pi n}\right)^{1/4} [1 + O(n^{-1}) + O(y^2 n^{-1})], \quad (1.11)$$

in which we suppose that y remains finite as $n \rightarrow \infty$. This may be obtained by using the known behavior of Hermite polynomials of large order.⁵

When $Q \gg \text{rms } I_n$ so that a is very large the distribution approaches that of a sine wave, namely

$$p_1(y) \sim \begin{cases} 0, & |y| > a \\ (a^2 - y^2)^{-1/2}/\pi, & |y| < a \end{cases} \quad (1.12)$$

$$\int_{-\infty}^y p_1(y_1) dy_1 \sim \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{y}{a}, \quad |y| < a$$

In order to study the manner in which the limiting expressions (1.12) are approached it is convenient to make the change of variable

$$\begin{aligned} x &= y - a \cos \theta, & d\theta &= [a^2 - (y - x)^2]^{-1/2} dx \\ z &= x - y + a \end{aligned}$$

in (1.6). We obtain

$$\begin{aligned} p_1(y) &= \frac{1}{\pi} \int_{y-a}^{y+a} \varphi(x) [a^2 - (y - x)^2]^{-1/2} dx \\ &= \frac{1}{\pi} \int_0^{2a} \varphi(z + y - a) [z(2a - z)]^{-1/2} dz. \end{aligned} \quad (1.13)$$

An asymptotic (as a becomes large) expression for $p_1(y)$ suitable for the middle portion of the distribution where $a - |y| \gg 1$ may be obtained from the first integral in (1.13). Since the principal contributions to the value of the integral come from the region around $x = 0$ we are led to expand the radical in powers of x and integrate termwise. Legendre polynomials enter naturally since they are sometimes defined as the coefficients in such an expansion. Replacing the limits of integration $y + a$ and $y - a$ by $+\infty$ and $-\infty$, respectively and integrating termwise gives

$$\begin{aligned} p_1(y) &\sim \frac{(a^2 - y^2)^{-1/2}}{\pi} \left[1 + \sum_{n=1}^{\infty} (-)^n \frac{1.3.5 \dots (2n-1)}{(a^2 - y^2)^n} P_{2n}(it^{1/2}) \right] \\ &= \frac{(a^2 - y^2)^{-1/2}}{\pi} \left[1 + \frac{3t + 1}{2(a^2 - y^2)} + \frac{3(35t^2 + 30t + 3)}{8(a^2 - y^2)^2} + \dots \right] \end{aligned} \quad (1.14)$$

⁵ A suitable asymptotic formula is given in *Orthogonal Polynomials*, by G. Szegő, *Am. Math. Soc. Colloquium, Pub.*, Vol. 23, (1939), p. 195.

where $t = y^2/(a^2 - y^2)$ and $P_{2n}(\)$ denotes the Legendre polynomial of order $2n$. We have written this as an asymptotic expansion because it obviously is one when y , and hence t , is zero in which case

$$P_{2n}(0) = (-)^n \frac{1.3.5 \cdots (2n-1)}{2.4 \cdots 2n}$$

When y is near a or is greater than a , a suitable asymptotic expansion may be obtained from the second integral in (1.13) by expanding $(2a - z)^{-1/2}$ in powers of $z/(2a)$ and integrating termwise. The upper limit of integration, $2a$, may be replaced by ∞ since $\varphi(z + y - a)$ may be assumed to be negligibly small when z exceeds $2a$. We thus obtain

$$\begin{aligned} p_1(y) &\sim \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n}{n!} \left(\frac{1}{2a}\right)^{n+1/2} \int_0^{\infty} \varphi(z + y - a) z^{n-1/2} dz \\ &= \frac{\varphi(y - a)}{\pi} \sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n}{n!} \left(\frac{1}{2a}\right)^{n+1/2} \int_0^{\infty} e^{-z(y-a)-(z^2/2)} z^{n-1/2} dz \end{aligned} \quad (1.15)$$

where we have used the notation

$$(\alpha)_0 = 1, \quad (\alpha)_n = \alpha(\alpha + 1) \cdots (\alpha + n - 1).$$

The integrals occurring in (1.15) are related to the parabolic cylinder function⁶ $D_m(x)$. Their properties may be obtained from the known properties of these functions or may be obtained by working directly with the integrals.

Suppose now that a is very large so that only the leading term in the series (1.15) for $p_1(y)$ need be retained.

Then

$$p_1(y) \sim a^{-1/2} F(y - a) \quad (1.16)$$

where

$$F(s) = \pi^{-1} 2^{-1/2} \int_0^{\infty} \varphi(z + s) z^{-1/2} dz \quad (1.17)$$

By writing out $\varphi(z + s)$, expanding $\exp(-zs)$ in a power series, and integrating termwise we see that

$$\begin{aligned} F(s) &= \frac{\varphi(s) 2^{-5/4}}{\pi} \sum_{\ell=0}^{\infty} \frac{\Gamma\left(\frac{\ell}{2} + \frac{1}{4}\right)}{\ell!} (-s\sqrt{2})^{\ell} \\ &= (2\pi)^{-1} s^{1/2} \varphi(s/\sqrt{2}) K_{\frac{1}{2}}(s^2/4) \end{aligned} \quad (1.18)$$

where K denotes a modified Bessel function.⁷ The relation (1.18) may also

⁶ Whittaker and Watson, "Modern Analysis," 4th ed. (1927), 347-351.

⁷ A table of $K_{\frac{1}{2}}(x)$ is given by H. Carsten and N. McKerrow, Phil. Mag. S7, Vol. 35 (1944), 812-818.

be obtained from pair 923.1 of "Fourier Integrals for Practical Applications," by G. A. Campbell and R. M. Foster.⁸

A curve showing $F(y-a)$ plotted as a function of $y-a$ is given in Fig. 3. It was obtained from the relation

$$F(s) = 2^{1/4} \pi^{-3/2} \chi(-s/\sqrt{2})$$

where

$$\chi(x) = \int_0^\infty e^{-(x-w^2)^2} dw$$

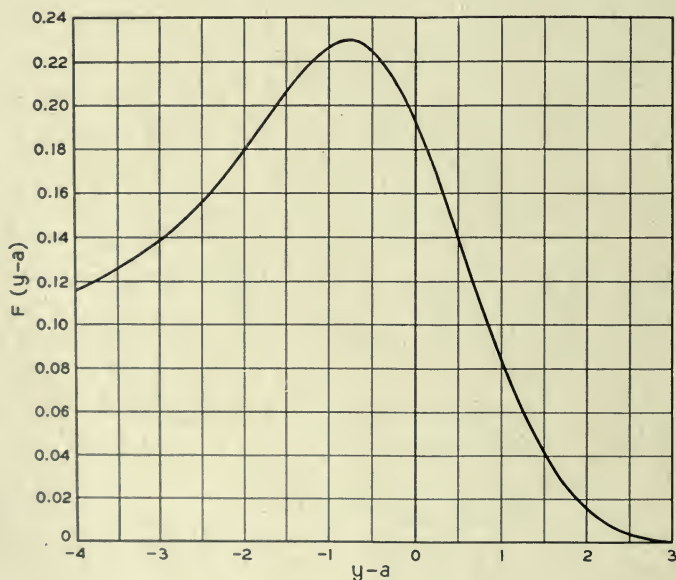


Fig. 3—Probability density of sine wave plus noise.

When $\text{rms } I_N \ll Q$ and I is near Q , $p_1(y) \sim a^{-1/2} F(y-a)$, $y-a = (I-Q)/(\text{rms } I_N)$.

See Fig. 1 for notation.

This function has been tabulated by Hartree and Johnston.⁹

The probability that I exceeds Q , or that y exceeds a , is, integrating the second of expressions (1.13),

$$\int_a^\infty p_1(y) dy = \frac{1}{\pi} \int_0^{2a} \frac{dz}{\sqrt{z(2a-z)}} \int_z^\infty \varphi(x) dx.$$

An asymptotic expansion may be obtained by expanding $(2a-z)^{-1/2}$ as in the derivation of (1.15) but we shall be content with the leading term.

⁸ Bell Telephone System Monograph B-584.

⁹ Manchester Lit. and Phil. Soc. Memoirs, v. 83, 183-188, Aug., 1939.

Using

$$\int_0^\infty z^{-1/2} dz \int_z^\infty \varphi(x) dz = \int_0^\infty \varphi(x) dx \int_0^x z^{-1/2} dz = 2^{1/4} \pi^{-1/2} \Gamma(\frac{3}{4})$$

we obtain

$$\int_a^\infty p_1(y) dy \sim 2^{-1/4} \pi^{-3/2} \Gamma(\frac{3}{4}) a^{-1/2} = 0.185 \dots a^{-1/2} \quad (1.19)$$

For use in computations we list the following values

$$\Gamma(\frac{1}{4}) = 3.62561, \quad \Gamma(\frac{3}{4}) = 1.22542, \quad \Gamma(\frac{5}{4}) = 0.90640$$

2. EXPECTED NUMBER OF CROSSINGS OF I PER SECOND

In this section, we shall study two questions. First, what is the probability $P(I_1, t_1)dt$ of I increasing through the value I_1 (i.e. of I passing through the value I_1 with positive slope) during the infinitesimal interval $t_1, t_1 + dt$? Second, what is the expected number $N(I_1)$ of times per second I increases through the value I_1 . When I_1 is zero, $2N(0)$ is the expected number of zeros per second, and when I_1 is large $N(I_1)$ is approximately equal to the expected number of maxima lying above the value I_1 in an interval one second long.

We start on the first question by considering the random function

$$z = F(a_1, a_2, \dots, a_N; t)$$

where the a 's are random variables. The probability that the random curve obtained by plotting z as a function of t increases through the value $z = z_1$ in the interval $t_1, t_1 + dt$ is¹⁰

$$dt \int_0^\infty \eta p(z_1, \eta; t_1) d\eta \quad (2.1)$$

where $p(\xi, \eta; t_1)$ denotes the probability density of the random variables

$$\xi = F(a_1, a_2, \dots, a_N; t_1)$$

$$\eta = \left[\frac{\partial F}{\partial t} \right]_{t=t_1}.$$

In our case z becomes the current I defined by equation (1.1). The method used to obtain equation (3.3-9) of Reference A may also be used to show that the quantity $p(I_1, \eta, t_1)$ (which now appears in (2.1)) is given by

$$p(I_1, \eta, t_1) = \frac{\pi N_0}{-\psi_0'} \varphi(y - a \cos qt_1) \varphi(x + b \sin qt_1) \quad (2.2)$$

¹⁰ This result is a straightforward generalization of expression (3.3-5) in Section 3.3 of Reference A where references to related results by M. Kac are given. A formula equivalent to (2.1) has also been given by Mr. H. Bondi in an unpublished paper written in 1944. He applies his formula to the problem studied in Section 4.

where $\varphi(\cdot)$ denotes the normal law function defined by equation (1.3) and

$$-\psi_0'' = 4\pi^2 \int_0^\infty w(f) f^2 df, \quad N_0 = \frac{1}{\pi} \sqrt{\frac{-\psi_0''}{\psi_0}}, \quad y = \frac{I_1}{\sqrt{\psi_0}}, \quad (2.3)$$

$$a = \frac{Q}{\sqrt{\psi_0}}, \quad x = \frac{\eta}{\sqrt{-\psi_0''}}, \quad b = \frac{Qq}{\sqrt{-\psi_0''}} = \frac{2af_q}{N_0}.$$

Equation (3.3-11) of Reference A shows that N_0 is the expected number of zeros per second which I_N would have if it were to flow alone.

Let $P(I_1, t_1)dt$ be the probability that I will increase through the value I_1 during the interval $t_1, t_1 + dt$. Then (2.1) and (2.2) give

$$P(I_1, t_1) = \int_0^\infty \eta p(I_1, \eta, t_1) d\eta \quad (2.4)$$

$$= \pi N_0 \varphi(y - a \cos qt_1) \int_0^\infty x \varphi(x + b \sin qt_1) dx.$$

The integral in (2.4) is of the form

$$\begin{aligned} \int_0^\infty x \varphi(x + v) dx &= \varphi(v) - v \int_v^\infty \varphi(x) dx \\ &= -\frac{v}{2} + \varphi(v) + v \int_0^v \varphi(x) dx \\ &= -v + \varphi(v) + v\varphi_{-1}(v) \\ &= -\frac{v}{2} + (2\pi)^{-1/2} {}_1F_1\left(-\frac{1}{2}; \frac{1}{2}; -\frac{v^2}{2}\right) \end{aligned} \quad (2.5)$$

where v replaces $b \sin qt_1$ and ${}_1F_1$ denotes a confluent hypergeometric function.

The distribution of the crossings at various portions of the cycle (of the sine wave) may be obtained by giving special values to qt_1 in (2.4).

The expected number of times I increases through the value I_1 in one second is

$$\begin{aligned} N(I_1) &= \text{Limit}_{T \rightarrow \infty} \frac{1}{T} \int_0^T P(I_1, t_1) dt_1 \\ &= N_0 \int_0^\pi \varphi(y - a \cos \theta) \left[\varphi(b \sin \theta) + b \sin \theta \int_0^{b \sin \theta} \varphi(x) dx \right] d\theta \end{aligned} \quad (2.6)$$

where we have used (2.4) and the second equation of (2.5). The integrand in (2.6) is composed of tabulated functions and is of a form suited to numerical integration. Expanding $\varphi(y - a \cos \theta)$ in (2.6) as in the derivation

of (1.10), replacing the quantity within the brackets by the series shown in the last equation of (2.5) and integrating termwise leads to

$$N(I_1) = N_0(\pi/2)^{1/2} \sum_{n=0}^{\infty} \frac{\varphi^{(2n)}(y)}{n!n!} \left(\frac{a}{2}\right)^{2n} {}_1F_1\left(-\frac{1}{2}; n+1; -\frac{b^2}{2}\right) \quad (2.7)$$

The series (2.7) converges for all values of a , y , and b . This follows from the inequality (1.11) which may be applied to $\varphi^{(2n)}(y)$, and from the fact that the ${}_1F_1$ is less than $\exp(b^2/2)$ as may be seen by comparing corresponding terms in their expansions.

The expected number of zeros, per second, of I is $2N(0)$ where we have set I_1 , and hence y , equal to zero. In this case the integral in (2.6) may be simplified somewhat and we obtain

$$2N(0) = N_0 \left[e^{-\alpha} I_0(\beta) + \frac{b^2}{2\alpha} Ie\left(\frac{\beta}{\alpha}, \alpha\right) \right] \quad (2.8)$$

where $I_0(\beta)$ is the Bessel function of order zero and imaginary argument and

$$\alpha = \frac{a^2 + b^2}{4}, \quad \beta = \frac{a^2 - b^2}{4}$$

$$Ie(k, x) = \int_0^x e^{-u} I_0(ku) du.$$

The integral $Ie(k, x)$ is tabulated in Appendix I.

I have been unable to obtain a simple derivation of (2.8). It was originally obtained from the following integral

$$N(I_1) = \frac{N_0}{2} \int_{-\pi}^{\pi} d\theta \varphi(y - a \cos \theta) \int_0^{\infty} x \varphi(x + b \sin \theta) dx \quad (2.9)$$

which may be derived from the second equation of (2.4) and the first of (2.6). Setting I_1 and y equal to zero and writing out the φ 's gives

$$2N(0) = \frac{N_0}{2\pi} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dx \\ x \exp \left[-\frac{1}{2}(x^2 + 2bx \sin \theta + a^2 \cos^2 \theta + b^2 \sin^2 \theta) \right].$$

Equation (2.8) was obtained by applying the method of Appendix III to this expression.

3. DEFINITIONS AND SIMPLE PROPERTIES OF R AND Θ

The remaining portion of this paper is concerned with the envelope R and the corresponding phase angle θ . These quantities are introduced and some of their simpler properties discussed in this section.

Suppose that the frequency band associated with I_N is relatively narrow

and contains the sine wave frequency f_q . The noise current may be resolved into two components, one "in phase" and the other "in quadrature" with $Q \cos qt$. Using the representation (2.8-6) of reference A and proceeding as in Section 3.7 of that paper:

$$I_N = \sum_{n=1}^M c_n \cos (\omega_n t - \varphi_n) \quad (3.1)$$

$$\begin{aligned} &= \sum_{n=1}^M c_n \cos [(\omega_n - q)t - \varphi_n + qt] \\ &= I_c \cos qt - I_s \sin qt \end{aligned} \quad (3.2)$$

where

$$I_c = \sum_{n=1}^M c_n \cos [(\omega_n - q)t - \varphi_n] \quad (3.3)$$

$$I_s = \sum_{n=1}^M c_n \sin [(\omega_n - q)t - \varphi_n]$$

$$\omega_n = 2\pi f_n, \quad f_n = n\Delta f, \quad c_n^2 = 2w(f_n)\Delta f$$

$w(f)$ denotes the power spectrum of I_N and the φ_n 's are random variables distributed uniformly over the interval $(0, 2\pi)$.

The total current I may be written as

$$\begin{aligned} I &= Q \cos qt + I_N \\ &= (Q + I_c) \cos qt - I_s \sin qt \\ &= R \cos \theta \cos qt - R \sin \theta \sin qt \\ &= R \cos (qt + \theta) \end{aligned} \quad (3.4)$$

where we have introduced the envelope function R and the phase angle θ by means of

$$\begin{aligned} R \cos \theta &= Q + I_c \\ R \sin \theta &= I_s \end{aligned} \quad (3.5)$$

Since I_c and I_s are functions of t whose variations are relatively slow in comparison with those of $\cos qt$, the same is true of R and (usually) θ .

A graphical illustration of equations (3.4) and (3.5) which is often used is shown in Fig. 4.

In accordance with the usual convention used in alternating current theory, the vector OQ is supposed to be rotating about the origin O with angular velocity q . If I_N happened to have the frequency $q/2\pi$, its vector representation QT would be fixed relative to OQ . In general, however, the

length and inclination of QT will change due to the random fluctuations of I_N . Thus the point T will wander around on the plane of the figure. If rms I_N is much less than Q , T will be close to the point Q most of the time. In this case

$$\begin{aligned} R &= [(Q + I_c)^2 + I_s^2]^{1/2} \sim Q + I_c \\ \theta &= \tan^{-1} \frac{I_s}{Q + I_c} \sim \frac{I_s}{Q} \\ \frac{d\theta}{dt} &\sim \frac{d}{dt} \frac{I_s}{Q} = \frac{I'_s}{Q} \end{aligned} \quad (3.6)$$

and a number of statistical properties of R and θ may be obtained from the corresponding properties of noise alone when we note that I_c , I_s , and I_N behave like noise currents whose power spectra are concentrated in the lower portion of the frequency spectrum.

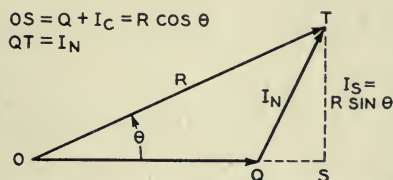


Fig. 4—Graphical representation of $I = Q \cos qt + I_N$.

By squaring both sides of equations (3.1) and (3.3) and then averaging with respect to t and the φ_n 's we may show that I_c , I_s , and I_N all have the same rms value, namely $\psi_0^{1/2}$.

It may be seen from (3.3) that the power spectra of I_c and I_s are both given by

$$w(f_q + f) + w(f_q - f) \quad (3.7)$$

where it is assumed that $0 \leq f \ll f_q$. Likewise the power spectrum of the time derivative I'_s of I_s is

$$4\pi^2 f^2 [w(f_q + f) + w(f_q - f)] \quad (3.8)$$

This follows from the representation of I'_s obtained by differentiating the expression (3.3) for I_s with respect to t , the procedure being the same as in the derivation of equation (7.2) in Section 7. The power spectra shown in Table 1 were computed from equations (3.7) and (3.8).

The correlation function for I_c , and hence also for I_s , is, from equations (A2-1) and (A2-3) of Appendix II,

$$\overline{I_c(t)I_c(t + \tau)} = g = \int_0^\infty w(f) \cos 2\pi(f - f_q)\tau df$$

where the bar denotes an average with respect to t and g is a function of τ . From (A2-3) the correlation function for I'_s is $-g''$ where the double prime on g denotes the second derivative with respect to τ .

Attention is sometimes fixed upon the variation in distance between successive zeros of I . The time between two successive zeros of I at, say, t_0 and t_1 is the time taken for $qt + \theta$, as appearing in $R \cos (qt + \theta)$, to increase by π . This assumes that the envelope R does not vanish in the interval. For the moment we write θ as $\theta(t)$ in order to indicate its dependence on the time t . Then t_0 and t_1 must satisfy the relation

$$qt_1 + \theta(t_1) - qt_0 - \theta(t_0) = \pi \quad (3.9)$$

Since $\theta(t)$ is a relatively slowly varying function we write

$$\theta(t_1) - \theta(t_0) = (t_1 - t_0)\theta'(t_0) + (t_1 - t_0)^2\theta''(t_0)/2 + \dots$$

TABLE 1
POWER SPECTRA OF I_N , I_c , I_s , AND I'_s

I_N	I_c and I_s	I'_s
$w(f) = w_0 = \psi_0/\beta$ for $f_q - \beta/2 < f < f_q + \beta/2$ $w(f) = 0$ elsewhere f_q = mid-band frequency	$2w_0$ for $0 < f < \beta/2$ 0 elsewhere	$8\pi^2 f^2 w_0$ for $0 < f < \beta/2$ 0 elsewhere
$w(f) = w_0 = \psi_0/\beta$ for $f_q - \beta < f < f_q$ $w(f) = 0$ elsewhere f_q = top frequency	w_0 for $0 < f < \beta$ 0 elsewhere	$4\pi^2 f^2 w_0$ for $0 < f < \beta$ 0 elsewhere
$w(f) = \frac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/(2\sigma^2)}$	$\frac{2\psi_0}{\sigma\sqrt{2\pi}} e^{-f^2/(2\sigma^2)}$	$\frac{8\pi^2 f^2 \psi_0}{\sigma\sqrt{2\pi}} e^{-f^2/(2\sigma^2)}$

where the primes denote differentiation with respect to t . When this is placed in (3.9) and terms of order $(t_1 - t_0)^2$ neglected, we obtain

$$\frac{1}{2(t_1 - t_0)} = \frac{q}{2\pi} + \frac{1}{2\pi} \theta'(t_0) \quad (3.10)$$

which relates the interval between successive zeros to θ' .

The expression on the right hand side of (3.10) may be defined as the instantaneous frequency:

$$\text{Instantaneous frequency} = f_q + \frac{1}{2\pi} \frac{d\theta}{dt} \quad (3.11)$$

This definition is suggested when $\cos 2\pi ft$ is compared with $\cos (qt + \theta)$ and also by (3.10) when we note that the period of the instantaneous fre-

quency is approximately equal to twice the distance between two successive zeros which is $2(t_1 - t_0)$.

The probability density of R is¹¹

$$\frac{R}{\psi_0} \exp \left[-\frac{R^2 + Q^2}{2\psi_0} \right] I_0(RQ/\psi_0) \quad (3.12)$$

where $I_0(RQ/\psi_0)$ denotes the Bessel function of order zero with imaginary argument. In Section 3.10 of Reference A, it is shown that the average value of R^n is*

$$\overline{R^n} = (2\psi_0)^{n/2} \Gamma\left(\frac{n}{2} + 1\right) {}_1F_1\left(-\frac{n}{2}; 1; -\rho\right), \quad (3.13)$$

where $\rho = Q^2/(2\psi_0)$, of which special cases are

$$\begin{aligned} \overline{R} &= e^{-\rho/2} (\pi\psi_0/2)^{1/2} [(1 + \rho)I_0(\rho/2) + \rho I_1(\rho/2)] \\ \overline{R^2} &= Q^2 + 2\psi_0 \end{aligned} \quad (3.14)$$

Curves showing the distribution of R are also given there.

4. EXPECTED NUMBER OF CROSSINGS OF R PER SECOND

Here we shall obtain expressions for the expected number N_R of times, per second, the envelope passes through the value R with positive slope. When R is large, N_R is approximately equal to the expected number of maxima of the envelope per second exceeding R and when R is small N_R is approximately equal to the expected number of minima less than R . For the special case in which the noise band is symmetrical and is centered on the sine wave frequency f_q N_R is given by the relatively simple expression (4.8).

The probability that the envelope passes through the value R during the interval $t, t + dt$ with positive slope is, from (2.1),

$$dt \int_0^\infty R' p(R, R', t) dR' \quad (4.1)$$

where $p(R, R', t)$ denotes the probability density of R and its time derivative R', t being regarded as a parameter.

An expression for $p(R, R', t)$ may be obtained from the probability density of I_c, I_s, I'_c, I'_s . From our representation of a noise current and the central limit theorem it may be shown (as is done for similar cases in Part III of Reference A) that the probability distribution of these four variables is

¹¹ In equation (60-A) of an unpublished appendix to his paper appearing in the *B.S.T.J.* Vol. 12 (1933), 35-75, Ray S. Hoyt gives an integral, obtained by integrating (3.12) with respect to R , for the cumulative distribution of R .

*The correlation function for the envelope of a signal plus noise, together with associated probability densities of the envelope and phase, is given by D. Middleton in a paper appearing soon in the *Quart. Jl. of Appl. Math.*

normal in four dimensions. If the variables be taken in the order given above the moment matrix is, from equations (A2-2) of Appendix II,

$$M = \begin{bmatrix} b_0 & 0 & 0 & b_1 \\ 0 & b_0 & -b_1 & 0 \\ 0 & -b_1 & b_2 & 0 \\ b_1 & 0 & 0 & b_2 \end{bmatrix} \quad (4.2)$$

where the b 's are defined by the integrals in equations (A2-1). The inverse matrix is

$$M^{-1} = \frac{1}{B} \begin{bmatrix} b_2 & 0 & 0 & -b_1 \\ 0 & b_2 & b_1 & 0 \\ 0 & b_1 & b_0 & 0 \\ -b_1 & 0 & 0 & b_0 \end{bmatrix}, \quad B = b_0 b_2 - b_1^2 \quad (4.3)$$

which may be readily verified by matrix multiplication, and the determinant $|M|$ is B^2 . The normal distribution may be written down at once when use is made of the formulas given in Section 2.9 of Reference A. The substitutions

$$\begin{aligned} I_c &= R \cos \theta - Q, & I'_c &= R' \cos \theta - R \sin \theta \theta' \\ I_s &= R \sin \theta, & I'_s &= R' \sin \theta + R \cos \theta \theta' \end{aligned} \quad (4.4)$$

$$dI_c dI_s dI'_c dI'_s = R^2 dR dR' d\theta d\theta'$$

enable us to write

$$\begin{aligned} &b_2(I_c^2 + I_s^2) + b_0(I_c'^2 + I_s'^2) \\ &\quad - 2b_1(I_c I_s' - I_s I_c') = b_2(R^2 - 2QR \cos \theta + Q^2) \\ &\quad \quad \quad + b_0(R'^2 + R^2 \theta'^2) \\ &\quad \quad \quad - 2b_1 R^2 \theta' + 2b_1 Q(R' \sin \theta + R \theta' \cos \theta). \end{aligned}$$

Consequently the probability density of R, R', θ, θ' is

$$\begin{aligned} p(R, R', \theta, \theta') &= \frac{R^2}{4\pi^2 B} \exp \left\{ -\frac{1}{2B} [b_2(R^2 - 2QR \cos \theta + Q^2) \right. \\ &\quad \left. + b_0(R'^2 + R^2 \theta'^2) - 2b_1 R^2 \theta' + 2b_1 Q(R' \sin \theta + R \theta' \cos \theta)] \right\} \end{aligned} \quad (4.5)$$

In this expression R ranges from 0 to ∞ , θ from $-\pi$ to π , and R' and θ' from $-\infty$ to $+\infty$. The probability density for R and R' is obtained by

integrating (4.5) with respect to θ and θ' over their respective ranges. The integration with respect to θ' may be performed at once giving

$$p(R, R', t) = \frac{R(2\pi)^{-3/2}}{\sqrt{Bb_0}} \int_{-\pi}^{\pi} d\theta \exp \left\{ -\frac{1}{2Bb_0} [B(R^2 - 2RQ\cos \theta + Q^2) + (b_0R' + b_1Q\sin \theta)^2] \right\} \quad (4.6)$$

From (4.1) and (4.6) it follows that the expected number N_R of times per second the envelope passes through R with positive slope is

$$N_R = \frac{R(2\pi)^{-3/2}}{\sqrt{Bb_0}} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} R' dR' \exp \left\{ -\frac{1}{2Bb_0} [B(R^2 - 2RQ\cos \theta + Q^2) + (b_0R' + b_1Q\sin \theta)^2] \right\} \quad (4.7)$$

When the power spectrum $w(f)$ of the noise current I_N is symmetrical about the sine wave frequency f_0 , b_1 is zero and B is equal to b_0b_2 . In this case the integrations in (4.7) may be performed. We obtain

$$\begin{aligned} N_R &= \left(\frac{b_2}{2\pi} \right)^{1/2} \frac{R}{b_0} I_0 \left(\frac{RQ}{b_0} \right) \exp \left(-\frac{R^2 + Q^2}{2b_0} \right) \\ &= \left(\frac{b_2}{2\pi} \right)^{1/2} \times \left[\text{Probability density of} \right. \\ &\quad \left. \text{envelope at the value } R \right] \end{aligned} \quad (4.8)$$

where the second line is obtained from expression (3.12). As will be seen from its definition (A2-1), b_0 is equal to the mean square value ψ_0 of I_N (and also of I_e and I_s).

Introducing the notation

$$\begin{aligned} v &= Rb_0^{-1/2} = R/\text{rms } I_N \\ a &= Ab_0^{-1/2} = Q/\text{rms } I_N, \end{aligned} \quad (4.9)$$

which is the same as that of equations (3.10-15) of Reference A except that there P denotes the amplitude of the sine wave and plays the same role as Q does here, enables us to write (4.8) as

$$N_R = \left[\frac{b_2}{2\pi b_0} \right]^{1/2} v I_0(av) e^{-(v^2+a^2)/2} = \left[\frac{b_2}{2\pi b_0} \right]^{1/2} p(v). \quad (4.10)$$

The function $p(v)$ is plotted as a function of v for various values of a in Fig. 6, Section 3.10, of Reference A.

It is interesting to note that

$$(b_2/b_0)^{1/2}/\pi = \text{Expected number of zeros per second of } I_e \text{ (or of } I_s) \quad (4.11)$$

This relation, which is true even if the noise band is not symmetrical about f_q , follows from equation (3.3-11) of Reference A.

When $Q \gg \text{rms } I_N$ and f_q is not at the center of the noise band it is easier to obtain the asymptotic form of N_R from the approximation (3.6),

$$R \sim Q + I_c,$$

instead of the double integral (4.7). When $Q \gg \text{rms } I_N$ and R is in the neighborhood of Q (as it is most of the time in this case), N_R is approximately equal to the expected number of times I_c increases through the value $I_{c1} = R - Q$ in one second. Thus, regarding I_c as a random noise current we have from expression (3.3-14) of Reference A

$N_R \sim e^{-I_{c1}^2/(2b_0)} \times [1/2 \text{ the expected number of zeros per second of } I_c]$ and when we use equation (4.11) we obtain

$$N_R \sim \frac{1}{2\pi} (b_2/b_0)^{1/2} e^{-(R-Q)^2/(2b_0)} = \frac{1}{2\pi} (b_2/b_0)^{1/2} e^{-(v-a)^2/2} \quad (4.12)$$

TABLE 2
 $w(f) = w_0 = b_0/\beta$ OVER A BAND OF WIDTH β

	b_2	N_R
1. Band extends from $f_q - \beta/2$ to $f_q + \beta/2$	$\pi^2 \beta^2 b_0/3$	$(\pi/6)^{1/2} \beta p(v) = 0.724 \beta p(v)$
2. Same as 1 and in addition $Q = 0$	"	$(\pi/6)^{1/2} \beta v e^{-v^2/2}$
3. Same as 1 and in addition $Q \gg \text{rms } I_N$	"	$\sim \frac{\beta}{2\sqrt{3}} e^{-(v-a)^2/2}$
4. Band extends from f_q to $f_q + \beta$ and $Q \gg \text{rms } I_N$	$4\pi^2 \beta^2 b_0/3$	$\sim \frac{\beta}{\sqrt{3}} e^{-(v-a)^2/2}$

Table 2 lists the forms assumed by (4.10) and (4.12) when the power spectrum $w(f)$ of the noise current I_N is constant over a frequency band of width β . The quantity b_0 in the expressions for b_2 represents the mean square value of I_N .

In the general case where the band of noise is not centered on f_q and where R is not large enough to make (4.12) valid we are obliged to return to the double integral (4.7). Some simplification is possible, but not as much as could be desired. Introducing the notation

$$\alpha = RQ/b_0, \quad \gamma = b_1 Q (Bb_0)^{-1/2}$$

$$x = (b_0 R' + b_1 Q \sin \theta) (Bb_0)^{-1/2}$$

enables us to write (4.7) as

$$N_R = (2\pi)^{-3/2} (R/b_0)(B/b_0)^{1/2} e^{-(R^2+Q^2)/(2b_0)} \int_{-\pi}^{\pi} d\theta \int_{\gamma \sin \theta}^{\infty} (x - \gamma \sin \theta) e^{\alpha \cos \theta - x^2/2} dx \quad (4.13)$$

Part of the integrand may be integrated with respect to x and the remaining portion integrated by parts with respect to θ . The double integral in the second line of (4.13) then becomes

$$\begin{aligned} & \int_{-\pi}^{\pi} e^{\alpha \cos \theta - (\gamma \sin \theta)^2/2} d\theta + \int_{-\pi}^{\pi} \left[\int_{\gamma \sin \theta}^{\infty} e^{-x^2/2} dx \right] d[\gamma \alpha^{-1} e^{\alpha \cos \theta}] \\ &= \int_{-\pi}^{\pi} (1 + \gamma^2 \alpha^{-1} \cos \theta) e^{\alpha \cos \theta - (\gamma \sin \theta)^2/2} d\theta \\ &= \gamma \alpha^{-1} e^{-(\gamma^2 + \alpha^2 \gamma^{-2})/2} \int_{-\pi}^{\pi} (\gamma \cos \theta + \alpha/\gamma) e^{(\gamma \cos \theta + \alpha/\gamma)^2/2} d\theta \\ &= 2\pi \sum_{n=0}^{\infty} \frac{(\frac{1}{2})_n}{n!} \left(-\frac{\gamma^2}{\alpha} \right)^n [I_n(\alpha) + \gamma^2 \alpha^{-1} I_{n+1}(\alpha)]. \end{aligned} \quad (4.14)$$

The series is obtained by expanding $\exp [-(\gamma \sin \theta)^2/2]$ in the second equation in powers of $\sin \theta$ and integrating termwise.

5. PROBABILITY DENSITY OF $\frac{d\theta}{dt}$

As was pointed out in Section 3 the time derivative θ' of the phase angle θ associated with the envelope is closely related to the instantaneous frequency. The probability density $p(\theta')$ of θ' may be expressed in terms of modified Bessel functions as shown by equation (5.4). Curves for $p(\theta')$ are given when the sine wave frequency f_q lies at the middle of a symmetrical band of noise. Although the expressions for $p(\theta')$ are rather complicated, those for the averages $\bar{\theta}'$ and $|\bar{\theta}'|$ given by equations (5.7) and (5.16) are relatively simple.

The probability density $p(\theta')$ may be obtained by integrating the expression (4.5) for $p(R, R', \theta, \theta')$ with respect to R, R', θ . The integration with respect to R' , the limits being $-\infty$ and $+\infty$, gives the probability density for R, θ, θ' :

$$p(R, \theta, \theta') = \frac{R^2}{4\pi^2} \left(\frac{2\pi}{b_0 B} \right)^{1/2} \exp [-aR^2 + 2bR \cos \theta + c \sin^2 \theta - b_2 Q^2/(2B)] \quad (5.1)$$

where

$$\begin{aligned} B &= b_0 b_2 - b_1^2 & b &= Q(b_2 - b_1 \theta') / (2B) \\ a &= (b_2 - 2b_1 \theta' + b_0 \theta'^2) / (2B) & c &= Q^2 b_1^2 / (2B b_0) = b_1^2 \rho / B \\ \rho &= Q^2 / (2b_0) & \gamma &= b^2 / a \end{aligned} \quad (5.2)$$

and b_0, b_1, b_2 are given in Appendix II.

Integrating with respect to R gives the probability density for θ, θ' . Expanding $\exp(2bR \cos \theta)$ in powers of R and integrating termwise,

$$p(\theta, \theta') = \frac{1}{16\pi a} \left(\frac{2}{ab_0 B} \right)^{1/2} e^{c \sin^2 \theta - b_2 b_0 \rho / B} \sum_{n=0}^{\infty} \frac{n+1}{\Gamma\left(\frac{n}{2} + 1\right)} \left(\frac{b \cos \theta}{a^{1/2}} \right)^n \quad (5.3)$$

When we integrate θ from $-\pi$ to π to obtain $p(\theta')$ the terms for which n is odd disappear and we have to deal with the series, writing γ for b^2/a ,

$$\sum_{m=0}^{\infty} \frac{2m+1}{m!} (\gamma \cos^2 \theta)^m = (2\gamma \cos^2 \theta + 1) \exp(\gamma \cos^2 \theta)$$

Thus, the probability density of θ' is

$$\begin{aligned} p(\theta') &= \frac{1}{16\pi a} \left(\frac{2}{ab_0 B} \right)^{1/2} \int_{-\pi}^{\pi} (2\gamma \cos^2 \theta + 1) e^{c \sin^2 \theta + \gamma \cos^2 \theta - b_2 b_0 \rho / B} d\theta \\ &= \frac{1}{8a} \left(\frac{2}{ab_0 B} \right)^{1/2} \left[(\gamma + 1) I_0 \left(\frac{\gamma - c}{2} \right) + \gamma I_1 \left(\frac{\gamma - c}{2} \right) \right] \\ &\quad \exp \left[\frac{c + \gamma}{2} - \frac{b_2 b_0 \rho}{B} \right] \end{aligned} \quad (5.4)$$

From (5.2)

$$\begin{aligned} \gamma - c &= \rho \frac{b_2 - 2b_1 \theta'}{b_2 - 2b_1 \theta' + b_0 \theta'^2} \\ \frac{c + \gamma}{2} - \frac{b_2 b_0 \rho}{B} &= -\frac{\rho}{2} \frac{b_2 - 2b_1 \theta' + 2b_0 \theta'^2}{b_2 - 2b_1 \theta' + b_0 \theta'^2} \end{aligned} \quad (5.5)$$

It will be noted that for large values of $|\theta'|$ the probability density of θ' varies as $|\theta'|^{-3}$. Although this makes the mean square value of θ' infinite, the average values $\bar{\theta}'$ and $|\bar{\theta}'|$ of θ' and $|\theta'|$ still exist. In order to obtain $\bar{\theta}'$ it is convenient to return to (4.5) and write

$$\bar{\theta}' = \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR \int_{-\infty}^{+\infty} dR' \int_{-\infty}^{\infty} d\theta' \theta' p(R, R', \theta, \theta') \quad (5.6)$$

The integration with respect to θ' may be performed by setting $R\theta'$ equal to x and using

$$\int_{-\infty}^{+\infty} x e^{-\alpha x^2 + 2\beta x} dx = (\beta/\alpha)(\pi/\alpha)^{1/2} e^{\beta^2/\alpha}$$

The integral in R' reduces to a similar integral except that the factor x in the integrand is absent. Performing these two integrations and using the definition of B leads to

$$\begin{aligned} \bar{\theta}' = \frac{1}{2\pi} \frac{b_1}{b_0^2} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR (R - Q \cos \theta) \\ \exp \left[-\frac{1}{2b_0} (R^2 - 2QR \cos \theta + Q^2) \right] \end{aligned}$$

We may integrate at once with respect to R . When this is done $\cos \theta$ disappears and the integration with respect to θ becomes easy. Thus

$$\bar{\theta}' = (b_1/b_0) \exp [-Q^2/(2b_0)] = (b_1/b_0)e^{-\rho} \quad (5.7)$$

When the noise power spectrum is equal to w_0 in the band extending from $f_0 - \beta/2$ to $f_0 + \beta/2$ and is zero outside the band, $b_1 = 2\pi(f_0 - f_q)b_0$. Hence, from (3.11),

$$\begin{aligned} \text{ave. instantaneous frequency} = f_q + \bar{\theta}'/(2\pi) \\ = f_0 + (f_q - f_0)(1 - e^{-\rho}) \end{aligned} \quad (5.8)$$

In the remainder of this section we assume the power spectrum of the noise current to be symmetrical about the sine wave frequency f_q . In this case b_1 and c are zero, B is equal to b_0b_2 and (5.4) becomes

$$\begin{aligned} p(\theta') &= \frac{1}{2}(b_0/b_2)^{1/2}(1 + z^2)^{-3/2}e^{-\rho+y/2} \\ &\quad [(y+1)I_0(y/2) + yI_1(y/2)] \\ &= \frac{1}{2}(b_0/b_2)^{1/2}(1 + z^2)^{-3/2}e^{-\rho} {}_1F_1\left(\frac{3}{2}; 1; y\right) \end{aligned} \quad (5.9)$$

where ${}_1F_1$ denotes a confluent hypergeometric function¹² and

$$z^2 = b_0\theta'^2/b_2, \quad y = (\gamma)_{b_1=0} = \rho/(1 + z^2) \quad (5.10)$$

When the noise power spectrum is constant in the band extending from $f_q - \beta/2$ to $f_q + \beta/2$ (see Table 2, Section 4)

$$(b_2/b_0)^{1/2} = 3^{-1/2}\beta\pi, \quad z = 3^{1/2}\theta'/(\beta\pi) \quad (5.11)$$

¹² The relation used above follows from equation (66) (with misprint corrected) of W. R. Bennett's paper cited in connection with equation (1.2).

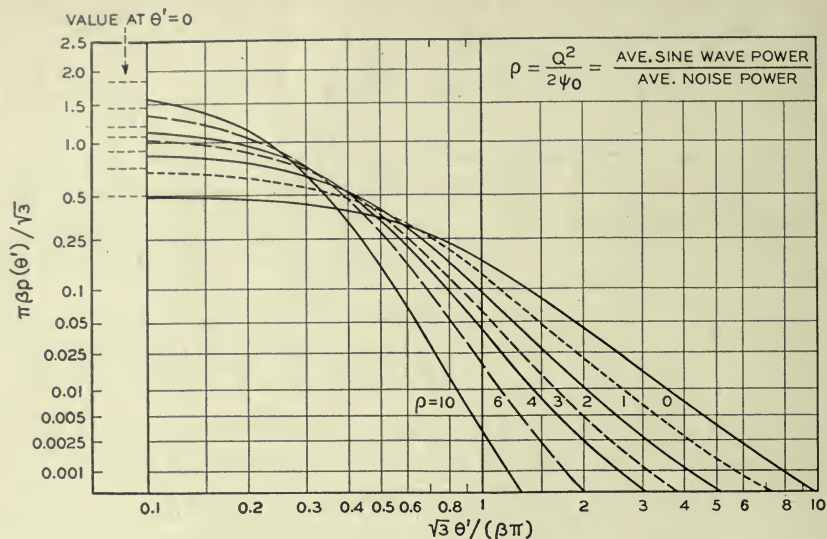


Fig. 5—Probability density of time derivative of phase angle.

$p(\theta') d\theta' =$ probability that the value of $d\theta/dt$ at an instant selected at random lies between θ' and $\theta' + d\theta'$. The power spectrum of I_N is constant in band of width β centered on f_q and is zero outside this band.

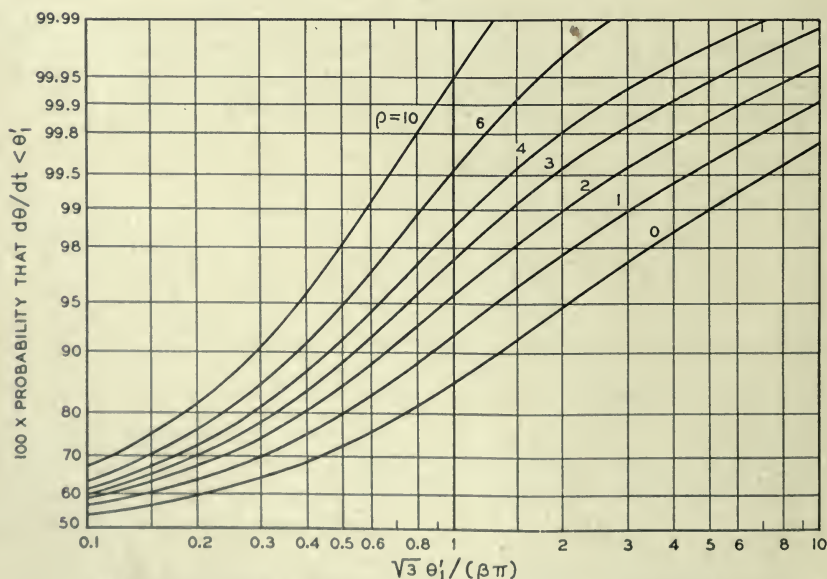


Fig. 6—Cumulative distribution of time derivative of phase angle.

Notation explained in Fig. 5.

The probability density $p(\theta')$ of θ' and its cumulative distribution, obtained by numerical integration, are shown in Figs. 5 and 6.

The probability that θ' exceeds a given θ'_1 is equal to the probability that z exceeds z_1 , where z_1 denotes $(b_0/b_2)^{1/2}\theta'_1$, and both probabilities are equal to

$$\frac{e^{-\rho}}{2} \int_{z_1}^{\infty} (1+z^2)^{-3/2} {}_1F_1 \left[\frac{3}{2}; 1; \rho(1+z^2)^{-1} \right] dz \quad (5.12)$$

The probability that $\theta' > \theta'_1$ becomes $e^{-\rho}/(4z_1^2)$ as $\theta'_1 \rightarrow \infty$.

When $Q \gg \text{rms } I_N$ the leading term in the asymptotic expansion of the ${}_1F_1$ in (5.9) gives

$$p(\theta') \sim \frac{1}{\sigma \sqrt{2\pi}} e^{-\theta'^2/(2\sigma^2)}, \quad \sigma^2 = b_2/Q^2 \quad (5.13)$$

when it is assumed that $z^2 \ll 1$. This expression holds only for the central portion of the curve for $p(\theta')$. Far out on the curve, $p(\theta')$ still varies as θ'^{-3} . Equation (5.13) may be obtained directly by using the approximation (3.6) that θ' is nearly equal to I'_s/Q and noticing that b_2 is the mean square value of I'_s .

If the sine wave is absent, ρ is zero and

$$p(\theta') = \frac{1}{2} (b_0/b_2)^{1/2} (1+z^2)^{-3/2} \quad (5.14)$$

which is consistent with the results given between equations (3.4-10) and (3.4-11) of Reference A. In this case (5.12) becomes

$$\frac{1}{2} - \frac{z_1}{2} (1+z_1^2)^{-1/2} \quad (5.15)$$

Although the standard deviation of θ' is infinite an idea of the spread of the distribution may be obtained from the average value of $|\theta'|$. Setting b_1 equal to zero in (4.5) in order to obtain the case in which the noise band is symmetrical about the sine-wave frequency leads to

$$\begin{aligned} |\overline{\theta'}| &= \frac{2}{4\pi^2 b_0 b_2} \int_0^{\infty} dR \int_{-\pi}^{\pi} d\theta \int_{-\infty}^{+\infty} dR' \int_0^{\infty} d\theta' \theta' R^2 \\ &\quad \exp \frac{1}{2} [-(R^2 - 2QR \cos \theta + Q^2)/b_0 - (R'^2 + R^2 \theta'^2)/b_2] \end{aligned}$$

The integrals in R' , θ' cause no difficulty and the integral in θ is proportional to the Bessel function $I_0(QR/b_0)$. When the resulting integral in R is evaluated¹³ we obtain

$$|\overline{\theta'}| = (b_2/b_0)^{1/2} e^{-\rho/2} I_0(\rho/2) \quad (5.16)$$

where $\rho = Q^2/(2b_0)$.

¹³ See, for example, G. N. Watson, "Theory of Bessel Functions," Cambridge (1944), p. 394, equation (5).

When ρ is zero equation (5.16) agrees with a result given in Section 3.4 of reference A, namely, for an ideal band pass filter

$$\frac{\text{ave } |\tau - \tau_1|}{\tau_1} = \frac{f_b - f_a}{\sqrt{3}(f_b + f_a)}$$

where τ is the interval between two successive zeros and τ_1 is its average value. τ is equal to $t_1 - t_0$ of our equation (3.10) from which it follows that

$$(\tau - \tau_1)/\tau_1 \approx -\theta'/q \quad (5.17)$$

6. EXPECTED NUMBER OF CROSSINGS OF θ AND $\frac{d\theta}{dt}$ PER SECOND

After a brief study of the expected number of times per second the phase angle θ increases through 0 and through π (where it is assumed that $-\pi < \theta \leq \pi$) expressions are obtained for the expected number $N_{\theta'}$ of times per second the time derivative of θ increases through the value θ' .

The point T shown in Fig. 4 of Section 3 wanders around, as time goes by, in the plane of the figure. How many times may we expect it to cross some preassigned section of the line OQ in one second? To answer this problem we note that, from expression (2.1), the probability that θ increases through zero during the interval $t, t + dt$ with the envelope lying between R and $R + dR$ is

$$dt dR \int_0^\infty \theta' p(R, 0, \theta') d\theta' \quad (6.1)$$

where the probability density in the integrand is obtained by setting θ equal to zero in equation (5.1). The expected number of such crossings per second is

$$(2\pi)^{-3/2} (b_0 B)^{-1/2} R^2 dR e^{-b_2(R-Q)^2/(2B)} \int_0^\infty d\theta' \theta' \exp [-b_0 R^2 \theta'^2/(2B) + b_1 R(R-Q)\theta'/B] \quad (6.2)$$

which may be evaluated in terms of error functions or the function $\varphi_{-1}(x)$ defined by equation (1.8). For the special case in which the power spectrum of the noise current I_N is symmetrical about the sine wave frequency, b_1 is zero and (6.2) yields

$$(2\pi)^{-3/2} b_0^{-1} b_2^{1/2} e^{-(R-Q)^2/(2b_0)} dR \quad (6.3)$$

From equation (6.1) onwards we have tacitly assumed that the range of θ is given by $-\pi < \theta \leq \pi$ because setting θ equal to any multiple of 2π in our equations leads to the same result as setting θ equal to zero. This is due to θ occurring only in $\cos \theta$ and $\sin \theta$. When θ increases through the value π ,

as it does when the point T crosses, in the downward direction, the extension of the line OQ lying to the left of the point O in Fig. 4, we imagine the value of θ to change discontinuously to the value $-\pi$.

The expected number of times per second θ increases through π may be obtained from (6.2) and, in the symmetrical case, from (6.3) by changing the sign of Q since this produces the same effect as changing θ from 0 to π in $p(R, \theta, \theta')$.

The expected number of crossings per second when R lies between two assigned values may be obtained by integrating the above equations. For example, the number of times per second θ increases through zero with R between Q and R_1 is, from (6.3) for the symmetrical case,

$$(4\pi)^{-1}(b_2/b_0)^{1/2} \operatorname{erf} [(2b_0)^{-1/2} | R_1 - Q |] \quad (6.4)$$

where we have used the absolute value sign to indicate that R_1 may be either less than or greater than Q and

$$\operatorname{erf} x = 2\pi^{-1/2} \int_0^x e^{-t^2} dt \quad (6.5)$$

Expressions for b_0 and b_2 are given by equations (A2-1) of Appendix II. The mean square value of I_N is b_0 , and when the power spectrum of I_N is constant over a band of width β , $b_2 = \pi^2 \beta^2 b_0 / 3$.

In much the same way it may be shown that the expected number of times per second θ increases through π with R between 0 and R_1 is

$$(4\pi)^{-1}(b_2/b_0)^{1/2} \{ \operatorname{erf} [(2b_0)^{-1/2}(R_1 + Q)] - \operatorname{erf} [(2b_0)^{-1/2}Q] \} \quad (6.6)$$

A check on these equations may be obtained by noting that the expected number of zeros per second of I_s , given by equation (4.11), is equal to twice the number of times θ increases through zero plus twice the number of times θ increases through π . Setting R_1 equal to zero in (6.4), to infinity in both (6.4) and (6.6), and adding the three quantities obtained gives half of (4.11), as it should.

Now we shall consider the crossings of θ' . The equations in the first part of the analysis are quite similar to those encountered in Section 3.8 of Reference A where the maxima of R , for noise alone, are discussed. We start by introducing the variables $x_1, x_2, \dots, x_6$ where

$$x_1 = I_c = R \cos \theta - Q, \quad x_4 = I_s = R \sin \theta \quad (6.7)$$

and the remaining x 's are defined in terms of the derivatives of I_c and I_s and are given by the equations just below (3.8-4) of Reference A.

Here we shall consider the noise band to be symmetrical about the sine

wave frequency f_q so that b_1 and b_3 are zero. Then from equations (3.8-3) and (3.8-4) of Reference A the probability density of $x_1, x_2, \dots, x_6$ is

$$\frac{1}{8\pi^3 b_2 D} \exp \left(-\frac{1}{2D} [b_4(x_1^2 + x_4^2) + 2b_2(x_1 x_3 + x_4 x_6) + (D/b_2)(x_2^2 + x_5^2) + b_0(x_3^2 + x_6^2)] \right) \quad (6.8)$$

where $D = b_0 b_4 - b_2^2$ and the b_n 's are given by equations (A2-1). Replacing the x 's by their expressions in terms of R and θ , similar to those just above equation (3.8-5) of Reference A, shows that the probability density for $R, R', R'', \theta, \theta', \theta''$ is

$$\begin{aligned} p(R, R', R'', \theta, \theta', \theta'') = & \frac{R^3}{8\pi^3 b_2 D} \exp \left(-\frac{1}{2D} [b_4(R^2 - 2RQ \cos \theta + Q^2) \right. \\ & + (D/b_2)(R'^2 + R^2 \theta'^2) + 2b_2(RR'' - R^2 \theta'^2) \\ & + b_0(R''^2 - 2RR'' \theta'^2 + 4R'^2 \theta'^2 + 4RR' \theta' \theta'' + R^2 \theta'^4 + R^2 \theta''^2) \\ & \left. - 2b_2 Q(R'' \cos \theta - R \theta'^2 \cos \theta - 2R' \theta' \sin \theta - R \theta'' \sin \theta) \right] \end{aligned} \quad (6.9)$$

It must be remembered that (6.9) applies only to the symmetrical case in which b_1 and b_3 are zero.

Integrating R' and R'' in (6.9) from $-\infty$ to ∞ gives the probability density of $R, \theta, \theta', \theta''$. The integration with respect to R'' is simplified by changing to the variable $R'' - R \theta'^2$. The result is

$$\begin{aligned} p(R, \theta, \theta', \theta'') = & R^3 (2\pi)^{-2} (b_0 b_2 D)^{-1/2} (1 + u)^{-1/2} \\ & \exp \left(-\frac{1}{2b_0} \left[R^2 - 2RQ \cos \theta + Q^2 + b_0 R^2 \theta'^2 / b_2 \right. \right. \\ & \left. \left. + \frac{(Q b_2 \sin \theta + b_0 R \theta'')^2}{(1 + u) D} \right] \right) \end{aligned} \quad (6.10)$$

where $u = 4b_2 b_0 \theta'^2 / D$. The expected number of times per second the time derivative of θ increases through the value θ' is

$$\begin{aligned} N_{\theta'} = & \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dR \int_0^{\infty} d\theta'' \theta'' p(R, \theta, \theta', \theta'') \\ = & \pi^{-2} (b_2 \delta / b_0)^{1/2} \int_{-\pi}^{\pi} d\theta \int_0^{\infty} r dr \int_0^{\infty} x dx \\ & \exp [-\gamma r^2 + 2r\alpha \cos \theta - \alpha^2 - \delta(x + \alpha \sin \theta)^2] \end{aligned} \quad (6.11)$$

where we have set

$$\begin{aligned} r &= R(2b_0)^{-1/2} & x &= rb_0\theta''/b_2 \\ \alpha &= Q(2b_0)^{-1/2} = \rho^{1/2} & \gamma &= 1 + b_0\theta'^2/b_2 = 1 + z^2 \\ \delta &= \frac{b_2^2}{(1+u)D} \end{aligned} \quad (6.12)$$

r being regarded as a constant when the variable of integration is changed from θ'' to x .

The double integral in θ and x occurring in (6.11) is of the same form as (A3-1) of Appendix III and hence may be transformed into (A3-3). Here $a = r\alpha$, $c = -\delta\alpha^2$, $c + b^2 = 0$. The diameter of the path of integration C may be chosen so large that the order of integration may be interchanged and the integration with respect to r performed. The result is again an integral of the form (A3-3) in which $a^2 = 0$. When this is reduced to (A3-6) it becomes

$$\begin{aligned} N_{\theta'} &= e^{-\rho}(2\pi\gamma)^{-1}b_2^{1/2}(b_0\delta)^{-1/2} [e^{-\delta\rho/2}I_0(\delta\rho/2) \\ &+ (1 + \gamma\delta)(1 + \gamma\delta/2)^{-1}e^{\rho/2}Ie\{\gamma\delta(2 + \gamma\delta)^{-1}, \rho/\gamma + \delta\rho/2\}] \end{aligned} \quad (6.13)$$

where we have used $Ie(-k, x) = Ie(k, x)$ which follows from the definition (A1-1) given in Appendix I.

When there is no sine wave present, ρ is zero and (6.13) becomes

$$N_{\theta'} = \frac{1}{2\pi\gamma} \left(\frac{b_2}{b_0\delta} \right)^{1/2} = \frac{\sqrt{\frac{b_4}{b_2} - \frac{b_2}{b_0} + 4\theta'^2}}{2\pi \left(1 + \frac{b_0}{b_2} \theta'^2 \right)} \quad (6.14)$$

This gives a partial check on some of the above analysis since (6.14) may be obtained immediately by setting α equal to zero in (6.11). Another check may be obtained by letting $\rho \rightarrow \infty$ and using $Ie(k, \infty) = (1 - k^2)^{-1/2}$. (6.13) becomes

$$N_{\theta'} \sim (2\pi)^{-1}(b_4/b_2)^{1/2}e^{-\rho z^2} \quad (6.15)$$

which agrees with the result obtained from $\theta' \sim I'_s/Q$.

For the case in which the power spectrum $w(f)$ of the noise is equal to the constant value w_0 over the frequency band extending from $f_q - \beta/2$ to $f_q + \beta/2$,

$$b_0 = \beta w_0, \quad b_2 = \pi^2\beta^3 w_0/3 = \pi^2\beta^2 b_0/3, \quad b_4 = \pi^4\beta^5 w_0/5 = \pi^4\beta^4 b_0/5 \quad (6.16)$$

These lead to

$$z = (b_0/b_2)^{1/2}\theta' = 3^{1/2}\theta'/(\pi\beta) \quad D/b_2^2 = b_4b_0/b_2^2 - 1 = 9/5 - 1 = 4/5$$

$$u = 4b_2^2z^2/D = 5z^2 \quad \delta = \frac{5}{4(1+5z^2)} \quad (6.17)$$

$$\gamma = 1 + z^2$$

and the coefficient in (6.13) may be simplified by means of

$$\frac{1}{2\pi\gamma} \left(\frac{b_2}{b_0\delta} \right)^{1/2} = \frac{\beta}{1+z^2} \left(\frac{1+5z^2}{15} \right)^{1/2} \quad (6.18)$$

From (6.14) we see that (6.18) is equal to $N_{\theta'}$ when noise alone is present (and is of constant strength in the band of width β). The curves of $N_{\theta'}/\beta$ versus z shown in Fig. 7 were obtained by setting (6.17) and (6.18) in (6.13). $N_{\theta'}/\beta$ approaches $e^{-\rho}/(z\sqrt{3})$ as $z \rightarrow \infty$.

When the wandering point T of Fig. 4 passes close to the point O , θ changes rapidly by approximately π and produces a pulse in θ' . In discussions of frequency modulation θ' is sometimes regarded as a noise voltage which is applied to a low pass filter. Although the closer T comes to O the higher the pulse, the area under the pulse will be of the order of π and the response of the low pass filter may be calculated approximately.

That the pulses in θ' arise in the manner assumed above may be checked as follows. We choose a point relatively far out on the curve for $\rho = 5$ in Fig. 7, say $z = \sqrt{3}\theta'/(\pi\beta) = 1.6$ or $\theta' = 2.9\beta$. The number of pulses per second having peaks higher than 2.9β is roughly $N_{\theta'} = .009\beta$, and half of these have peaks greater than $\theta' = 3.8\beta$ which is obtained from Fig. 7 for $N_{\theta'} = .0045\beta$. From Fig. 6 we see that θ' exceeds 2.9β about .0018 of the time. Thus the average width at the height 2.9β of the class of pulses whose peaks exceed this value is $.0018/(\cdot 009\beta) = .2/\beta$ seconds. This figure is to be checked by the width obtained from the assumption that the typical pulse arises when T moves along a straight line with speed v and passes within a distance b of O . We take $\tan \theta = vt/b = \alpha t$ so that $\theta' = \alpha/(1 + \alpha^2 t^2)$. From this expression for θ' it follows that a pulse of peak height 3.8β (the median height) has a width of $.3/\beta$ seconds at $\theta' = 2.9\beta$. This agreement seems to be fairly good in view of the roughness of our work. A similar comparison may be made for $\rho = 0$ by using the limiting forms of (5.15) and (6.18). Here it is possible to compute the average width instead of estimating it from the median peak value. Exact agreement is obtained, both methods leading to an average width of $\pi/(4\theta')$ seconds at height θ' .

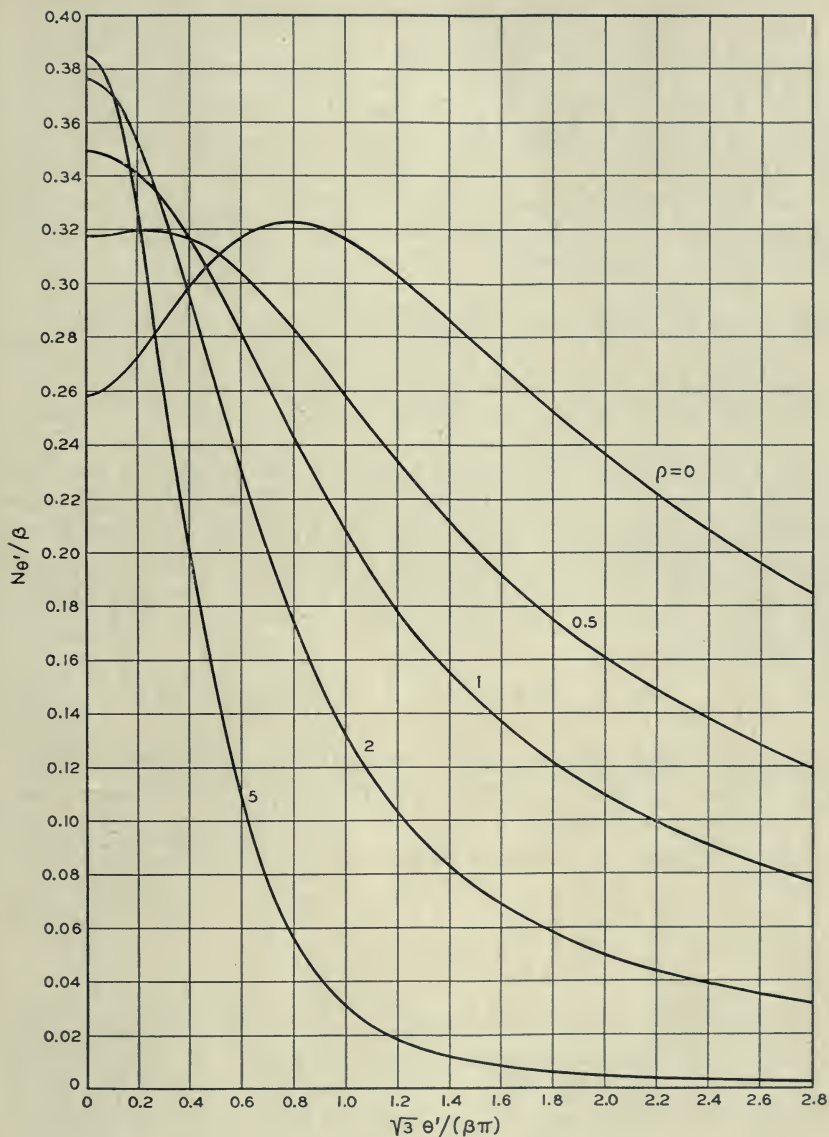


Fig. 7—Crossings of time derivative of phase angle.

$N\theta'$ = expected number of times per second $d\theta/dt$ increases through the value θ' . ρ , β , and the power spectrum of I_N are the same as in Fig. 5.

7. CORRELATION FUNCTION FOR $\frac{d\theta}{dt}$

In this section we shall compute the correlation function $\Omega(\tau)$ of $\theta'(t)$. We are primarily interested in $\Omega(\tau)$ because it is, according to a fundamental result due to Wiener, the Fourier transform¹⁴ of the power spectrum $W(f)$ of $\theta'(t)$.

We shall first consider the case in which the sine wave power is very large compared with the noise power so that, from (3.6), θ is approximately I_s/Q and θ' approximately I'_s/Q . Then using (A2-3) and (A2-1)

$$\begin{aligned}\Omega(\tau) &= \overline{\theta'(t)\theta'(t+\tau)} \approx Q^{-2} \overline{I'_s(t)I'_s(t+\tau)} \\ &= -g''Q^{-2} = 4\pi^2Q^{-2} \int_0^\infty w(f)(f-f_q)^2 \cos 2\pi(f-f_q)\tau df\end{aligned}\quad (7.1)$$

When $w(f)$ is effectively zero outside a relatively narrow band in the neighborhood of f_q , as it is in the cases with which we shall deal, (7.1) leads to the relation (divide the interval $(0, \infty)$ into $(0, f_q)$ and (f_q, ∞) , introduce new variables of integration $f_1 = f_q - f$, $f_2 = f - f_q$ in the respective intervals, replace the upper limit f_q of the first integral by ∞ , combine the integrals, and compare with (2.1-6) of Reference A)

Power spectrum of $\theta'(t) = W(f)$

$$\approx 4\pi^2f^2Q^{-2}[w(f_q + f) + w(f_q - f)] \quad (7.2)$$

This form is closely related to results customarily used in frequency modulation studies. It should be remembered that in (7.2) it is assumed that $0 < f \ll f_q$ and rms $I_N \ll Q$.

Additional terms in the approximation for $\Omega(\tau)$ may be obtained by expanding

$$\theta = \arctan \frac{I_s}{Q + I_c}$$

in descending powers of Q , multiplying two such series (one for time t and the other for time $t + \tau$) together, and averaging over t . If I_{c1} , I_{s1} and I_{c2} , I_{s2} denote the values of I_c , I_s at times t and $t + \tau$ respectively, the average values of the products of the I 's may be obtained by expanding the characteristic function (obtainable from equation (7.5) given below by setting $z_5 = z_6 = z_7 = z_8 = 0$) of the four random variables I_{c1} , I_{s1} , I_{c2} , I_{s2} . This method is explained in Section 4.10 of Reference A. When $w(f)$ is symmetrical about f_q it is found that

¹⁴ The form which we shall use is given by equation (2.1-5) of Reference A.

$$\begin{aligned}
 \overline{\theta_1 \theta_2} &= \frac{g}{Q^2} + \frac{g^2}{Q^4} + \frac{8g^3}{3Q^6} + \dots \\
 \Omega(\tau) &= \overline{\theta_1' \theta_2'} = -\frac{d^2}{d\tau^2} \overline{\theta_1 \theta_2} \\
 &= -\frac{g''}{Q^2} - \frac{2}{Q^4} (gg'' + g'^2) - \frac{8}{Q^6} (g^2 g'' + 2g'^2 g) + \dots
 \end{aligned} \tag{7.3}$$

From the exact expression for $\Omega(\tau)$ obtained below it is seen that the last equation in (7.3) is really asymptotic in character and the series does not converge. We infer that this is also true for the first equation of (7.3).

We shall now obtain the exact expression for the correlation function $\Omega(\tau)$ of $\theta'(t)$ when f_a is at the center of a symmetrical band of noise. At first sight it would appear that the easiest procedure is to calculate the correlation function for $\theta(t)$ and then obtain $\Omega(\tau)$ by differentiating twice. However, difficulties present themselves in getting θ outside the range $-\pi, \pi$ since θ enters the expressions only as the argument of trigonometrical functions. Because I could not see any way to overcome this difficulty I was forced to deal with θ' directly. Unfortunately this increases the complexity since now the distribution of the time derivatives of I_c and I_s also must be considered.

We have

$$\begin{aligned}
 \tan \theta &= \frac{I_s}{Q + I_c}, & \sec^2 \theta &= 1 + \left(\frac{I_s}{Q + I_c} \right)^2 \\
 \theta' &= \frac{(Q + I_c)I'_s - I_s I'_c}{\sec^2 \theta (Q + I_c)^2} = \frac{(Q + I_c)I'_s - I_s I'_c}{(Q + I_c)^2 + I_s^2}
 \end{aligned}$$

and the value of $\theta'(t)\theta'(t + \tau)$ is the eight-fold integral

$$\begin{aligned}
 \Omega(\tau) &= \int_{-\infty}^{+\infty} dI_{c1} \dots \int_{-\infty}^{+\infty} dI'_{s2} p(I_{c1}, \dots, I'_{s2}) \\
 &\quad \frac{(Q + I_{c1})I'_{s1} - I_{s1}I'_{c1}}{(Q + I_{c1})^2 + I_{s1}^2} \times \frac{(Q + I_{c2})I'_{s2} - I_{s2}I'_{c2}}{(Q + I_{c2})^2 + I_{s2}^2}
 \end{aligned} \tag{7.4}$$

where $p(I_{c1}, \dots, I'_{s2})$ is an eight-dimensional normal probability density. As before, the subscripts 1 and 2 refer to times t and $t + \tau$, respectively. The most direct way of evaluating the integral (7.4) is to insert the expression for $p(I_{c1}, \dots, I'_{s2})$ and then proceed with the integration. Indeed, this method was used the first time the integral (7.4) was evaluated. Later it was found that the algebra could be simplified by representing $p(I_{c1}, \dots, I'_{s2})$ as the Fourier transform of its characteristic function. The second procedure will be followed here.

The characteristic function for $I_{c1}, I_{c2}, I_{s1}, I_{s2}, I'_{c1}, I'_{c2}, I'_{s1}, I'_{s2}$ is, from (A2-2) and (A2-3) of Appendix II and Section 2.9 of Reference A,

$$\begin{aligned} \text{ave. exp } i[z_1 I_{c1} + z_2 I_{c2} + z_3 I_{s1} + z_4 I_{s2} + z_5 I'_{c1} + z_6 I'_{c2} + z_7 I'_{s1} + z_8 I'_{s2}] \\ = \exp \left(-\frac{1}{2} \right) [b_0(z_1^2 + z_2^2 + z_3^2 + z_4^2) + b_2(z_5^2 + z_6^2 + z_7^2 + z_8^2) \\ + 2b_1(z_1 z_7 + z_2 z_8 - z_3 z_5 - z_4 z_6) \\ + 2g(z_1 z_2 + z_3 z_4) + 2g'(z_1 z_6 - z_2 z_5 + z_3 z_8 - z_4 z_7) \\ - 2g''(z_5 z_6 + z_7 z_8) + 2h(z_1 z_4 - z_2 z_3) \\ + 2h'(z_1 z_8 + z_2 z_7 - z_3 z_6 - z_4 z_5) - 2h''(z_5 z_8 - z_6 z_7)]. \end{aligned} \quad (7.5)$$

Since we have included b_1, h, h', h'' this holds when f_q is not necessarily at the center of the noise band. However, henceforth we return to our assumption that f_q is placed at the center of a symmetrical noise band and take b_1, h, h', h'' to be zero.

The probability density of $I_{c1}, \dots, I'_{s2}$ which is to be placed in (7.4) is the eight-fold integral

$$\begin{aligned} p(I_{c1}, \dots, I'_{s2}) = (2\pi)^{-8} \int_{-\infty}^{+\infty} dz_1 \dots \int_{-\infty}^{+\infty} dz_8 \\ \exp [-iz_1 I_{c1} - \dots - iz_8 I'_{s2}] \times [\text{ch.f.}] \end{aligned} \quad (7.6)$$

where "ch.f." denotes the characteristic function obtained by setting b_1, h, h', h'' equal to zero on the right hand side of (7.5).

The integral (7.4) for $\Omega(\tau)$ may be written as

$$\Omega(\tau) = J_1 - J_2 - J_3 + J_4 \quad (7.7)$$

where J_1 is the 16-fold integral

$$\begin{aligned} J_1 = \int_{-\infty}^{+\infty} dI_{c1} \dots \int_{-\infty}^{+\infty} dI'_{s2} (2\pi)^{-8} \int_{-\infty}^{+\infty} dz_1 \dots \int_{-\infty}^{+\infty} dz_8 \\ \exp [-iz_1 I_{c1} - \dots - iz_8 I'_{s2}] \\ \frac{(Q + I_{c1})(Q + I_{c2})I'_{s1}I'_{s2}}{[(Q + I_{c1})^2 + I_{s1}^2][(Q + I_{c2})^2 + I_{s2}^2]} \times [\text{ch.f.}] \end{aligned} \quad (7.8)$$

and J_2, J_3, J_4 are obtained from J_1 by replacing the product $(Q + I_{c1})(Q + I_{c2})I'_{s1}I'_{s2}$ by $(Q + I_{c1})I'_{s1}I_{s2}I'_{c2}$, $I_{s1}I'_{c1}(Q + I_{c2})I'_{s2}$, $I_{s1}I'_{c1}I_{s2}I'_{c2}$ respectively.

The integration with respect to I_{c1} and I_{s1} in (7.8) may be performed at once. We replace $Q + I_{c1}$ and I_{s1} by x and y , respectively, and use

$$\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \frac{x}{x^2 + y^2} e^{-izx - i\zeta y} = \frac{-2\pi iz}{z^2 + \zeta^2}. \quad (7.9)$$

The integration with respect to I_{c2} and I_{s2} may be performed in a similar manner. In this way we obtain a 12-fold integral.

The integrations with respect to the I 's may be performed by using

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{+\infty} dI \int_{-\infty}^{+\infty} e^{-izI} f(z) dz &= f(0) \\ \frac{1}{2\pi} \int_{-\infty}^{+\infty} I dI \int_{-\infty}^{+\infty} e^{-izI} f(z) dz &= -i \left[\frac{df(z)}{dz} \right]_{z=0} \end{aligned} \quad (7.10)$$

The result is the four-fold integral

$$\begin{aligned} J_1 &= (2\pi)^{-2} \int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_4 \frac{z_1 z_2 (g'' - g'^2 z_3 z_4)}{(z_1^2 + z_3^2)(z_2^2 + z_4^2)} \\ &\exp [-(b_0/2)(z_1^2 + z_2^2 + z_3^2 + z_4^2) - g(z_1 z_2 + z_3 z_4) + iQ(z_1 + z_2)]. \end{aligned} \quad (7.11)$$

In the same way J_2, J_3, J_4 may be reduced to the integrals obtained from (7.11) by replacing $z_1 z_2 (g'' - g'^2 z_3 z_4)$ by $-g'^2 z_1^2 z_4^2, -g'^2 z_2^2 z_3^2$ and $z_3 z_4 (g'' - g'^2 z_1 z_2)$, respectively. When the J 's are combined in accordance with (7.7) we obtain an integral which may be obtained from (7.11) by replacing $z_1 z_2 (g'' - g'^2 z_3 z_4)$ by

$$g''(z_1 z_2 + z_3 z_4) + g'^2(z_1 z_4 - z_2 z_3)^2 \quad (7.12)$$

The terms $z_1^2 + z_3^2$ and $z_2^2 + z_4^2$ in the denominator may be represented as infinite integrals. Interchanging the order of integration and expressing (7.12) in terms of partial derivatives of an exponential function leads to the six-fold integral

$$\begin{aligned} \Omega(\tau) &= (4\pi)^{-2} \int_0^\infty du \int_0^\infty dv \left[-g'' \frac{\partial}{\partial g} + g'^2 \frac{\partial^2}{\partial \alpha^2} \right]_{\alpha=0} \int_{-\infty}^{+\infty} dz_1 \cdots \int_{-\infty}^{+\infty} dz_4 \\ &\exp [-(b_0 + u)(z_1^2 + z_3^2)/2 - (b_0 + v)(z_2^2 + z_4^2)/2 \\ &\quad - g(z_1 z_2 + z_3 z_4) - \alpha(z_1 z_4 - z_2 z_3) + iQ(z_1 + z_2)] \end{aligned} \quad (7.13)$$

where the subscript $\alpha = 0$ indicates that α is to be set equal to zero after the differentiations are performed.

When the four-fold integral in the z 's is evaluated (7.13) becomes

$$\begin{aligned} \Omega(\tau) &= \int_0^\infty du \int_0^\infty dv \left[-g'' \frac{\partial}{\partial g} + g'^2 \frac{\partial^2}{\partial \alpha^2} \right]_{\alpha=0} \\ &\quad \frac{1}{4D} \exp [-Q^2(2b_0 - 2g + u + v)/(2D)] \quad (7.14) \\ &= \int_0^\infty du \int_0^\infty dv [(g'^2 - gg'')(2 - 2F + Q^2/g) - g'^2 Q^2/g] e^{-F/(4D_0^2)} \end{aligned}$$

where

$D = (b_0 + u)(b_0 + v) - g^2 - \alpha^2$, $F = Q^2(2b_0 - 2g + u + v)/(2D_0)$ and D_0 denotes the value of D obtained by setting $\alpha = 0$. When differentiating with respect to α it is helpful to note that

$$\frac{\partial^2 f(D)}{\partial \alpha^2} = f''(D) \left(\frac{\partial D}{\partial \alpha} \right)^2 + f'(D) \frac{\partial^2 D}{\partial \alpha^2}$$

and that only $f'(D) = df/dD$ need be obtained since $\partial D/\partial \alpha$ vanishes when $\alpha = 0$.

In order to reduce the double integral to a single integral we make the change of variables

$$r = Q^2(b_0 + u - g)/(2D_0) \equiv \frac{Q^2(b_0 + u - g)}{2[(b_0 + u)(b_0 + v) - g^2]}$$

$$s = Q^2(b_0 + v - g)/(2D_0), \quad F = r + s \quad (7.15)$$

$$\partial(r, s)/\partial(u, v) = -rs/D_0, \quad 4srD_0 = Q^2[Q^2 - 2g(r + s)]$$

The limits of integration for r and s are obtained by noting that the points $(0, 0)$, $(\infty, 0)$, (∞, ∞) , $(0, \infty)$ in the (u, v) plane go into $(Q^2/(2b_0 + 2g), Q^2/(2b_0 + 2g))$, $(Q^2/(2b_0), 0)$, $(0, 0)$, $(0, Q^2/(2b_0))$, respectively, in the (r, s) plane. It may be verified that the region of integration in the (r, s) plane is the interior of the quadrilateral obtained by joining the above points by straight lines. Equation (7.14) may now be written as

$$\Omega(\tau) = \iint \left\{ \frac{(g'^2 - gg'')(2 - 2r - 2s + Q^2/g) - g'^2 Q^2/g}{Q^2[Q^2 - 2g(r + s)]} \right\} e^{-r-s} dr ds$$

$$= \frac{g'^2 - gg''}{2g^2} y_1 - \frac{g'^2}{2g^2} y_2 \quad (7.16)$$

where y_1 and y_2 are the dimensionless quantities

$$y_1 = \iint \frac{2g^2(2 - 2r - 2s + Q^2/g)}{Q^2[Q^2 - 2g(r + s)]} e^{-r-s} dr ds$$

$$y_2 = \iint \frac{2ge^{-r-s} dr ds}{Q^2 - 2g(r + s)}$$

It is seen that

$$y_1 = 2gQ^{-2} \left\{ y_2 + \iint e^{-r-s} dr ds \right\}. \quad (7.17)$$

Since the integrands are functions of $r + s$ alone we are led to apply the transformation

$$\iint_A f(r + s) dr ds = \int_0^\alpha u f(u) du + \int_\alpha^{2\beta} \frac{\alpha(2\beta - u)}{2\beta - \alpha} f(u) du \quad (7.18)$$

where A is the area enclosed by the quadrilateral whose vertices are at the points (r, s) given by $(0, 0)$, $(0, \alpha)$, (β, β) , $(\alpha, 0)$ and it is assumed that β and α are positive. u is a new variable and is not the one introduced in (7.13).

Setting $\alpha = Q^2/(2b_0)$ and $\beta = Q^2/(2b_0 + 2g)$, using (7.18), and introducing the notation

$$\begin{aligned} \rho &= Q^2/(2b_0), & k &= g/b_0 \\ \xi &= Q^2/(2g) = \rho/k, & \lambda &= \frac{Q^2}{b_0 + g} = \frac{2\rho}{1 + k} \end{aligned} \quad (7.19)$$

permits us to write

$$\begin{aligned} \iint_A e^{-r-s} dr ds &= \int_0^\rho u e^{-u} du + \int_\rho^\lambda \frac{\rho(\lambda - u)}{\lambda - \rho} e^{-u} du \\ &= 1 - \frac{\lambda e^{-\rho}}{\lambda - \rho} + \frac{\rho e^{-\lambda}}{\lambda - \rho} \end{aligned} \quad (7.20)$$

and (7.17) yields

$$y_2 = \frac{\rho}{k} y_1 - 1 + \frac{2}{1 - k} e^{-\rho} - \frac{1 + k}{1 - k} e^{-2\rho/(1+k)} \quad (7.21)$$

where we have expressed λ in terms of ρ and k .

The double integral defining y_2 may be treated in the same way as (7.20):

$$y_2 = \iint_A \frac{e^{-r-s} dr ds}{\xi - r - s} = \int_0^\rho \frac{u e^{-u} du}{\xi - u} + \int_\rho^\lambda \frac{\rho(\lambda - u) e^{-u} du}{(\lambda - \rho)(\xi - u)}$$

Writing $u = \xi - (\xi - u)$ and $\lambda - u = \lambda - \xi + (\xi - u)$ in the two numerators leads to

$$\begin{aligned} y_2 &= \xi \int_0^\rho \frac{e^{-u} du}{\xi - u} - \int_0^\rho e^{-u} du \\ &\quad - \xi \int_\rho^\lambda \frac{e^{-u} du}{\xi - u} + \frac{\rho}{\lambda - \rho} \int_\rho^\lambda e^{-u} du \end{aligned} \quad (7.22)$$

where we have used $\rho(\lambda - \xi)/(\lambda - \rho) = -\xi$ to simplify the coefficient of the third integral. When the second and fourth integrals are evaluated, their contribution to y_2 is found to be equal to the terms independent of y_1 on the right of (7.21). Hence, comparison of equations (7.21) and (7.22) shows that

$$y_1 = \int_0^\rho \frac{e^{-u} du}{\xi - u} - \int_\rho^\lambda \frac{e^{-u} du}{\xi - u} \quad (7.23)$$

The integrals in (7.23) may be evaluated in terms of the exponential integral $Ei(x)$ defined by, for x real,

$$Ei(x) = \int_{-\infty}^x e^t dt/t = C + \frac{1}{2} \log_e x^2 + \sum_{n=1}^{\infty} \frac{x^n}{n!n} \\ \sim e^x \sum_{n=0}^{\infty} n!/x^{n+1}$$

where $C = .577 \dots$ is Euler's constant and Cauchy's principal value of the integral is to be taken when $x > 0$. We set $t = \xi - u$ and obtain

$$y_1 = e^{-\rho/k} \left\{ Ei[\rho/k] - 2Ei[\rho(1-k)/k] + Ei \left[\frac{\rho(1-k)}{k(1+k)} \right] \right\}$$

where we have again expressed ξ and λ in terms of ρ and k .

A power series for y_1 which converges when $-1/3 \leq k < 1$ may be obtained by expanding the denominators of the integrands in (7.23) in powers of u/ξ and integrating termwise:

$$y_1 = \xi^{-1} [1 - 2e^{-\rho} + e^{-\lambda}] \\ + 1! \xi^{-2} [1 - 2(1 + \rho/1!)e^{-\rho} + (1 + \lambda/1!)e^{-\lambda}] \\ + 2! \xi^{-3} [1 - 2(1 + \rho/1! + \rho^2/2!)e^{-\rho} + (1 + \lambda/1! + \lambda^2/2!)e^{-\lambda}] \\ + \dots \quad (7.24)$$

The following special values may be obtained from the equation given above. When $\rho = 0$

$$y_1 = -\log_e (1 - k^2) \\ y_2 = 0 \quad (7.25)$$

This result may also be obtained by evaluating the integral obtained when we set $Q = 0$, $z_1 = r_1 \cos \theta_1$, $z_3 = r_1 \sin \theta_1$, $z_2 = r_2 \cos \theta_2$, $z_4 = r_2 \sin \theta_2$ in (7.11) and (7.12).

Near $k = 1$,

$$y_1 \approx e^{-\rho} [Ei(\rho) - C - \log_e \rho(1 - k^2)] \\ y_2 \approx \rho y_1 - 1 + (1 + \rho)e^{-\rho} \quad (7.26)$$

Near $k = 0$,

$$y_1 \approx k(1 - e^{-\rho})^2/\rho, \quad y_2 \approx y_1 \quad (7.27)$$

except when $\rho = 0$ in which case y_1 is approximately k^2 .

When ρ is large

$$y_1 \sim \frac{k}{\rho} + \frac{1!k^2}{\rho^2} + \frac{2!k^3}{\rho^3} + \frac{3!k^4}{\rho^4} + \dots \\ y_2 \sim -1 + \frac{\rho}{k} y_1 \sim \frac{1!k}{\rho} + \frac{2!k^2}{\rho^2} + \dots \quad (7.28)$$

except near $k = 1$ where both y_1 and y_2 have logarithmic infinities. The asymptotic expansion (7.3) for $\Omega(\tau)$, which was obtained by the first method

of this section, may be checked by inserting (7.28) in the expression (7.16) for $\Omega(\tau)$ in terms of y_1 and y_2 .

Values of y_1 and y_2 tabulated as functions of k for various values of ρ are given in Table 3. Negative values of k have not been considered since they

TABLE 3
VALUES OF y_1 AND y_2 USED IN COMPUTATION OF CORRELATION FUNCTION OF $d\theta/dt$
 $\Omega(\tau) = \overline{\theta'(t)\theta'(t + \tau)} = [g'^2(y_1 - y_2) - gg''y_1]/(2g^2)$
 $g \equiv g(\tau) = \int_0^\infty w(f) \cos 2\pi(f - f_q)\tau df, \quad k = g(\tau)/g(0)$

k	Values of y_1					Values of y_2			
	ρ					ρ			
	0	.5	1	2	5	.5	1	2	5
0	0	0	0	0	0	0	0	0	0
.1	.01005	.03526	.04224	.03854	.02000	.03171	.04147	.03936	.02051
.2	.04082	.08043	.09003	.07979	.04105	.06550	.08654	.08275	.04283
.3	.09431	.1379	.1452	.1246	.06292	.1022	.1363	.1315	.06702
.4	.1744	.2110	.2102	.1740	.08586	.1432	.1926	.1870	.09384
.5	.2877	.3056	.2886	.2296	.1101	.1914	.2579	.2515	.1238
.6	.4463	.4278	.3860	.2942	.1358	.2481	.3368	.3289	.1576
.7	.6733	.5953	.5129	.3721	.1636	.3220	.4379	.4269	.1975
.8	1.0216	.8416	.6914	.4729	.1941	.4275	.5803	.5602	.2461
.84	1.2228	.9798	.7888	.5242	.2075	.4866	.6593	.6318	.2693
.88	1.4890	1.1590	.9127	.5866	.2219	.5641	.7619	.7226	.2964
.90	1.6607	1.2742	.9898	.6241	.2296	.6138	.8260	.7752	.3114
.92	1.8734	1.4144	1.0834	.6686	.2378	.6753	.9058	.8486	.3294
.94	2.1507	1.5948	1.2024	.7217	.2466	.7550	1.0093	.9333	.3498
.96	2.5459	1.8486	1.3668	.7939	.2566	.8711	1.1558	1.0546	.3752
.97	2.8285	2.0251	1.4815	.8414	.2623	.9474	1.2605	1.1366	.3849
.98	3.2289	2.2762	1.6405	.9073	.2690	1.0704	1.4081	1.2548	.4119
.99	3.9170	2.7080	1.9066	1.0127	.2778	1.2773	1.6610	1.4505	.4429
.995	4.6072	3.1341	2.1721	1.1125	.2846	1.4838	1.9175	1.6416	.4705
.997	5.1175	3.4445	2.3622	1.1866	.2889	1.6367	2.1048	1.7859	.4893

are not required for the case in which I_N has a normal law power spectrum, the case discussed in the next section.

8. POWER SPECTRUM OF $\frac{d\theta}{dt}$ WHEN I_N HAS NORMAL LAW POWER SPECTRUM

The problem of computing the power spectrum $W(f)$ of $\theta'(t)$ appears to be a difficult one.* In order to obtain an answer without an excessive amount of work we have had to do two things which are rather restrictive. First, we confine our attention to the case in which the power spectrum $w(f)$ of

*Since the above was written the general f. m. problem has been studied by D. Middleton. He generalizes our (7.11) and (7.12), introduces polar coordinates, expands the integrand in powers of g , and integrates termwise. $W(f)$ then follows somewhat as in a.m. theory.

I_N is of the normal law type (our method could be applied to other types but g' and g'' would be more complicated functions of τ and Table 3 would have to be extended to negative values of k , if they should occur). Second, we resort to numerical integration to obtain a portion of $W(f)$. Because of the second item our results are either tabulated or are given as curves, shown in Figs. 8 and 9, except when $Q = 0$ (noise only) in which case the power spectrum of θ' is given by the series (8.7).

The power spectrum of I_N is assumed to be

$$w(f) = \frac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/(2\sigma^2)} \quad (8.1)$$

The mean square value of I_N is equal to that of a noise current whose power spectrum has the constant value of $\psi_0/(\sigma\sqrt{2\pi})$ over a band of width $f_b - f_a = \sigma\sqrt{2\pi} = \sigma 2.507$. The value of $w(f)$ is one quarter of its mid-band value at the points $f - f_q = \pm\sigma\sqrt{2\log_e 4} = \pm\sigma 1.665$ (the 6 db points) and the distance between these points is 3.330σ . Integration of (8.1) shows that the mean square value of I_N is ψ_0 in accordance with our customary notation. The mid-band value of $w(f)$ is $\psi_0/(\sigma\sqrt{2\pi})$.

Assuming $f_q \gg \sigma$ and evaluating the integrals (A2-1) of Appendix II defining b_0 and g gives

$$\begin{aligned} b_0 &= \psi_0, & g &= \psi_0 e^{-2(\pi\sigma\tau)^2} = \psi_0 e^{-u^2/2} \\ g'/g &= -uu' = -2\pi\sigma u, & g''/g &= -(2\pi\sigma)^2(1 - u^2) \\ \frac{g''g - g'^2}{g^2} &= -(2\pi\sigma)^2, & k &= g/b_0 = e^{-u^2/2} \end{aligned} \quad (8.2)$$

where we have set

$$u = 2\pi\sigma\tau, \quad u' = 2\pi\sigma \quad (8.3)$$

and the primes on g and u denote differentiation with respect to τ . The correlation function is accordingly, from (7.16).

$$\Omega(\tau) = 2\pi^2\sigma^2(\gamma_1 - u^2\gamma_2) \quad (8.4)$$

If $\theta'(\iota)$ be regarded as a noise current its power spectrum is

$$W(f) = 4 \int_0^\infty \Omega(\tau) \cos 2\pi f\tau \, d\tau \quad (8.5)$$

When noise alone is present, ρ is zero and (7.25) yields

$$\Omega(\tau) = -2\pi^2\sigma^2 \log_e (1 - k^2) = -2\pi^2\sigma^2 \log_e (1 - e^{-u^2}) \quad (8.6)$$

In this case the power spectrum is, from (8.3), (8.5), and (8.6),

$$\begin{aligned} W_N(f) &= -4\pi\sigma \int_0^\infty \cos(uf/\sigma) \log(1 - e^{-u^2}) du \\ &= 2\sigma\pi^{3/2} \sum_{n=1}^\infty n^{-3/2} e^{-f^2/(4n\sigma^2)}, \end{aligned} \quad (8.7)$$

the series being obtained by expanding the logarithm and integrating term-wise. When this equation was used for computation it was found convenient to apply the Euler summation formula to sum the terms in the series beyond the $(N - 1)$ st. Writing b for $f^2/(4\sigma^2)$, the series in (8.7) becomes

$$\begin{aligned} &1^{-3/2}e^{-b/1} + 2^{-3/2}e^{-b/2} + \dots (N - 1)^{-3/2}e^{-b/(N-1)} \\ &+ (\pi/b)^{1/2} \operatorname{erf}[(b/N)^{1/2}] + N^{-3/2}e^{-b/N} \left[\frac{1}{2} - \frac{1}{12N} \left(-\frac{3}{2} + \frac{b}{N} \right) \right. \\ &\left. + \frac{1}{720N^3} \left(-\frac{105}{8} + \frac{105}{4} \frac{b}{N} - \frac{21}{2} \frac{b^2}{N^2} + \frac{b^3}{N^3} \right) + \dots \right] \end{aligned} \quad (8.8)$$

When b is zero the sum¹⁵ of the series is 2.61237 The values for $\rho = 0$ in Table 4 were computed by taking $N = 12$ in (8.8). As $b \rightarrow \infty$ the dominant term in (8.8) is seen to be the one containing erf (choose N so that $b = N^{3/2}$). Hence as $f \rightarrow \infty$

$$W_N(f) \sim 4\pi^2\sigma^2/f. \quad (8.9)$$

When both noise and the sine wave are present it is convenient to split the power spectrum into three parts. The first part, $W_1(f)$, is proportional to $W_N(f)$, the power spectrum with noise alone. The second part $W_2(f)$ is proportional to the form $W(f)$ assumes when $\text{rms } I_N \ll Q$ and the third part $W_3(f)$ is of the nature of a correction term. This procedure is suggested when we subtract the leading terms in the expressions (7.26) and (7.27) (corresponding to $k = 1$ and $k = 0$, respectively) from y_1 . Likewise we subtract the leading term in y_2 , (7.27), at $k = 0$ but do not bother to do so at the end $k = 1$ because u^2y_2 approaches zero there. We therefore write

$$\begin{aligned} y_1 - u^2y_2 &= [y_1 + e^{-\rho} \log(1 - k^2) - k(1 - e^{-\rho})^2/\rho - u^2y_2 \\ &+ u^2k(1 - e^{-\rho})^2/\rho] - e^{-\rho} \log(1 - k^2) + (1 - u^2)k(1 - e^{-\rho})^2/\rho \\ &= Z(u) - e^{-\rho} \log(1 - k^2) - \frac{g''(2\pi\sigma)^{-2}}{b_0\rho} (1 - e^{-\rho})^2 \end{aligned} \quad (8.10)$$

¹⁵ "Theory and Application of Infinite Series," Knopp, (1928), page 561.

where $Z(u)$ denotes the function enclosed by the brackets in the first equation and the expressions for g''/g and k in (8.2) have been used in the replacement of $(1 - u^2)k$.

TABLE 4
VALUES OF $W_3(f)/(4\pi^2\sigma)$

$\frac{\delta f}{\sigma\pi}$	$\rho = 0$	0.5	1.0	2.0	5.0
0	0	-.03517	-.03891	-.02444	-.001948
1	0	-.03003	-.03196	-.01830	-.001814
2	0	-.01717	-.01486	-.003304	.004052
3	0	-.002436	.004014	.01252	.008225
4	0	.008757	.01730	.02244	.01027
6	0	.01478	.02157	.02167	.007665
8	0	.01018	.01366	.01237	.003505
10	0	.005768	.007378	.006201	.001437
12	0	.004027	.004463	.003552	.0006439

VALUES OF $W(f)/(4\pi^2\sigma)$

0	.7369	.4118	.2322	.07529	.003017
1	.7098	.4294	.2672	.1134	.02342
2	.6439	.4516	.3231	.1784	.05828
3	.5542	.4225	.3225	.1947	.06852
4	.4623	.3496	.2654	.1580	.01590
6	.3195	.2178	.1508	.07554	.01540
8	.2390	.1553	.1019	.04506	.005325
10	.1908	.1215	.07768	.03206	.002726
12	.1595	.1003	.06306	.02511	.001719

Inserting (8.10) in the expression (8.4) for $\Omega(\tau)$ and taking the Fourier transform (8.5) leads to

$$\begin{aligned}
 W(f) &= W_1(f) + W_2(f) + W_3(f) \\
 W_1(f) &= e^{-\rho} W_N(f) \\
 W_2(f) &= -\frac{(1 - e^{-\rho})^2}{2b_0\rho} \int_0^\infty g'' \cos 2\pi f\tau \, d\tau \\
 &= \frac{(1 - e^{-\rho})^2}{\rho} (2\pi f)^2 \frac{e^{-f^2/(2\sigma^2)}}{\sigma\sqrt{2\pi}} \\
 W_3(f) &= 4\pi\sigma \int_0^\infty Z(u) \cos(uf/\sigma) \, du
 \end{aligned} \tag{8.11}$$

In these equations $W_N(f)$ is obtained from (8.7), and $W_2(f)$ by two-fold integration by parts to reduce g'' to g then evaluating the integral obtained

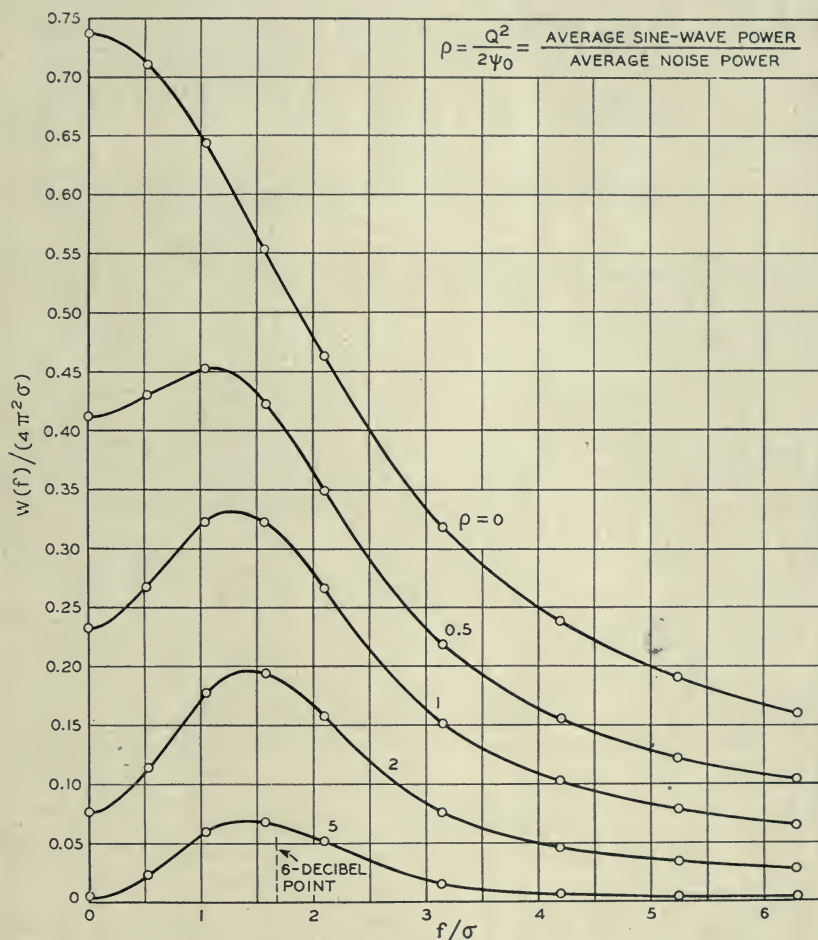


Fig. 8—Power spectrum of $d\theta/dt$.

Power spectrum of I_N is assumed to be

$$\psi_0(\sigma\sqrt{2\pi})^{-1} \exp [-(f - f_q)^2/(2\sigma^2)].$$

In this expression f is a frequency near f_q . The f in $W(f)$ and in the abscissa is a much lower frequency. $W(f)$ = power spectrum of $\theta' = d\theta/dt$, θ' being regarded as a random noise current. Dimensions of $W(f)df$ same as $(d\theta/dt)^2$ or $(\text{radians})^2/\text{sec}^2$.

by substituting the expression (8.2) for g . That $W(f)$ approaches $W_2(f)$ as $\rho \rightarrow \infty$ follows when expression (8.11) for $W_2(f)$ is compared with the limiting form (8.13) given below.

Instead of dealing with $W(f)$ it is more convenient to deal with $(4\pi^2\sigma)^{-1}W(f)$ which is the sum of the three components

$$\begin{aligned}(4\pi^2\sigma)^{-1}W_1(f) &= \frac{e^{-\rho}}{2\sqrt{\pi}} \sum_{n=1}^{\infty} n^{-3/2} e^{-f^2/(4n\sigma^2)} \\(4\pi^2\sigma)^{-1}W_2(f) &= \frac{(1 - e^{-\rho})^2}{\rho\sqrt{2\pi}} \left(\frac{f}{\sigma}\right)^2 e^{-f^2/(2\sigma^2)} \\(4\pi^2\sigma)^{-1}W_3(f) &= \frac{1}{\pi} \int_0^{\infty} Z(u) \cos(uf/\sigma) du\end{aligned}\quad (8.12)$$

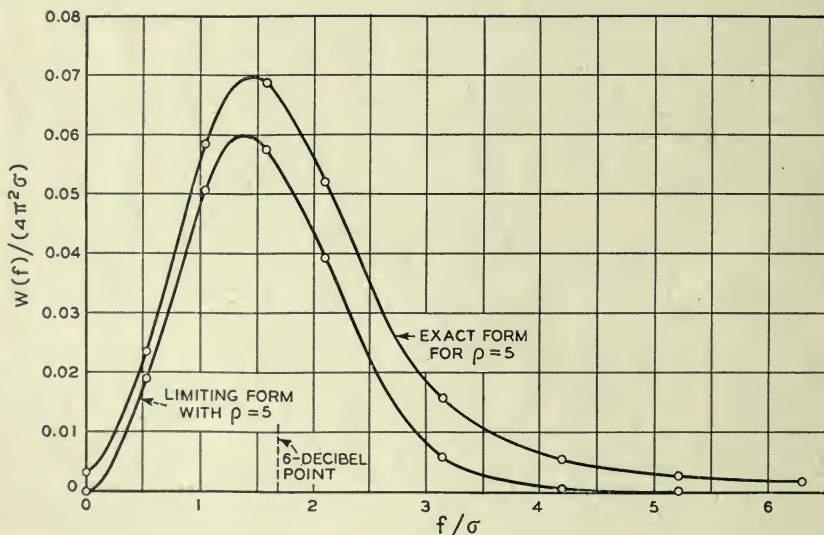


Fig. 9—Approach of $W(f)$ to limiting form.

As $\rho \rightarrow \infty$, $W(f) \rightarrow 4\pi^2\sigma (\rho\sqrt{2\pi})^{-1} (f/\sigma)^2 \exp[-f^2/(2\sigma^2)]$.

The integral involving $Z(u)$ has been computed by Simpson's rule, y_1 and y_2 being obtained from Table 3, with the results shown in the first section of Table 4. The value of $W_2(f)$ may be computed directly, and $W_1(f)$ may be obtained from $W_N(f)$. The values of these two functions together with those of $W_3(f)$ enable us to compute the values of $(4\pi^2\sigma)^{-1}W(f)$ given in Table 4 and plotted in Fig. 8.

Since, as is shown by (8.9), $W_N(f)$ varies as $1/f$ for large values of f , the areas under the curves of Fig. 8 become infinite. This agrees with the fact that the mean square value of θ' is infinite.

The values of $(4\pi^2\sigma)^{-1}W(0)$ for ρ equal to 0, .5, 1, 2, and 5 are .7369, 4118, 2322, .07529, and .003017 respectively. When these values are plotted on

semi-log paper they tend to lie on a straight line whose slope suggests that $W(0)$ decreases as $e^{-\rho}$ when ρ becomes large.

The limiting form assumed by $W(f)$ as $\rho \rightarrow \infty$ is given by equation (7.2). When the normal law expression (8.1) assumed in this section for the power spectrum of I_N is put in (7.2) we find that

$$W(f) \rightarrow \frac{4\pi^2\sigma}{\rho\sqrt{2\pi}} \left(\frac{f}{\sigma}\right)^2 e^{-f^2/(2\sigma^2)} \quad (8.13)$$

Fig. 9 shows that for $\rho = 5$ the limiting form (8.13) agrees quite well with the exact form computed above.

Both (7.2) and (8.13) show that, for small values of f , the power spectrum of θ' varies as f^2 when $\rho \gg 1$. This is in accord with Crosby's* result that the voltage spectrum of the random noise in the output of a frequency modulation receiver is triangular when the carrier to noise ratio is large. When this ratio becomes small he finds that the spectrum becomes rectangular. Fig. 8 shows this effect in that the areas under the curves between the ordinates at $f = 0$ and $f = \lambda\sigma$ (where λ is some number, generally less than unity, depending on the ratio of the widths of the i.f. and audio bands) become rectangles, approximately, as ρ decreases.

APPENDIX I

THE INTEGRAL $Ie(k, x)$

The integral¹⁶

$$Ie(k, x) = \int_0^x e^{-u} I_0(ku) du, \quad (A1-1)$$

where $I_0(ku)$ denotes the Bessel function of imaginary argument and order zero, occurs in Sections 2 and 6. The following special cases are of interest.

$$\begin{aligned} Ie(0, x) &= 1 - e^{-x} \\ Ie(1, x) &= xe^{-x}[I_0(x) + I_1(x)] \\ Ie(k, \infty) &= \frac{1}{\sqrt{1 - k^2}} \end{aligned} \quad (A1-2)$$

The second of these relations is due to Bennett.¹⁷

* M. G. Crosby, "Frequency Modulation Noise Characteristics," Proc. I. R. E. Vol. 25 (1937), 472-514. See also J. R. Carson and T. C. Fry, "Variable Electric Circuit Theory with Application to the Theory of Frequency Modulation," B.S.T.J. Vol. 16 (1937), 513-540.

¹⁶ The notation was chosen to agree with that used by Bateman and Archibald (Guide to Tables of Bessel Functions appearing in "Math. Tables and Aids to Comp.," Vol. 1 (1944) pp. 205-308) to discuss integrals used by Schwarz (page 248).

¹⁷ It is given in equation (62) of the reference cited in connection with our equation (1.2) in Section 1.

The values in the table given below were computed by Simpson's rule for numerical integration. The work was checked at several points by using

$$Ie(k, x) = \sum_{n=0}^{\infty} (k/2)^{2n} \frac{(2n)!}{n!n!} A_n$$

where

$$A_n = 1 - \left[1 + x + \frac{x^2}{2!} \cdots + \frac{x^{2n}}{(2n)!} \right] e^{-x}$$

When x is so large that $Ie(k, x)$ is nearly equal to $Ie(k, \infty)$ we have

$$Ie(k, x) \sim (1 - k^2)^{-1/2} - [2k(1 - k)]^{-1/2} (2/\sqrt{\pi}) \int_{t_1}^{\infty} e^{-t^2} dt$$

where $t_1 = \sqrt{x(1 - k)}$. However, this was not found to be especially useful in checking the values given in the table.

$$\text{TABLE OF } Ie(k, x) = \int_0^x e^{-u} I_0(ku) du$$

x	k						
	0	.2	.4	.6	.8	.9	1.0
0	0	0	0	0	0		0
.2	.1813	.1813	.1814	.1815	.1816		.1818
.4	.3297	.3298	.3303	.3311	.3322		.3337
.6	.4512	.4517	.4530	.4554	.4586		.4629
.8	.5507	.5516	.5545	.5593	.5661		.5749
1.0	.6321	.6337	.6386	.6468	.6584		.6736
.2	.6988	.7012	.7086	.7209	.7386		.7620
.4	.7534	.7567	.7669	.7841	.8089		.8422
.6	.7981	.8025	.8157	.8383	.8712		.9157
.8	.8347	.8401	.8566	.8850	.9267		.9839
2.0	.8647	.8712	.8910	.9255	.9766		1.0476
.2	.8892	.8968	.9201	.9607	1.0217		1.1075
.4	.9093	.9179	.9446	.9916	1.0627		1.1642
.6	.9257	.9354	.9655	1.0186	1.1001		1.2183
.8	.9392	.9499	.9831	1.0424	1.1345		1.2699
3.0	.9502	.9618	.9982	1.0635	1.1661		1.3195
.2	.9592	.9718	1.0110	1.0822	1.1953		1.3672
.4	.9666	.9800	1.0220	1.0988	1.2223		1.4132
.6	.9727	.9868	1.0314	1.1136	1.2475		1.4578
.8	.9776	.9925	1.0394	1.1268	1.2708		1.5010
4.0	.9817	.9971	1.0463	1.1386	1.2926		1.5430
.2	.9830	1.0010	1.0522	1.1492	1.3130		1.5839
.4	.9877	1.0043	1.0574	1.1587	1.3320		1.6237
.6	.9899	1.0070	1.0619	1.1672	1.3499		1.6625
.8	.9918	1.0092	1.0657	1.1749	1.3666		1.7005
5.0	.9933	1.0111	1.0690	1.1818	1.3823		1.7376
5.4	.9955	1.0140	1.0743	1.1937	1.4110		1.8095

TABLE—Continued

x	k						
	0	.2	.4	.6	.8	.9	1.0
5.8	.9970	1.0160	1.0783	1.2034	1.4364		1.8786
6.2	.9980	1.0174	1.0814	1.2114	1.4590		1.9452
6.6	.9986	1.0183	1.0837	1.2180	1.4792		2.0097
7.0	.9991	1.0190	1.0854	1.2234	1.4972		2.0722
7.4	.9994	1.0195	1.0867	1.2278	1.5134		2.1328
7.8	.9996	1.0198	1.0876	1.2375	1.5279		2.1917
8.2	.9997	1.0201	1.0885	1.2346	1.5409		2.2491
8.6	.9998	1.0202	1.0891	1.2371	1.5526		2.3050
9.0	.9999	1.0203	1.0896	1.2393	1.5631		2.3597
10.0	1.0000	1.0205	1.0902	1.2431	1.5852	1.9207	2.4910
11.0	1.0000	1.0206	1.0907	1.2456	1.6024	1.9668	2.6157
12.0	1.0000	1.0206	1.0909	1.2471	1.6158	2.0066	2.7347
13.0	1.0000	1.0206	1.0910	1.2482	1.6263	2.0411	2.8487
14.0	1.0000	1.0206	1.0910	1.2488	1.6346	2.0711	2.9584
15.0	1.0000	1.0206	1.0911	1.2492	1.6412	2.0973	3.0641
∞	1.0000	1.0206	1.0911	1.2500	1.6667	2.2942	∞

x	k			
	.86	.90	.96	1.0
15.0	1.8773	2.0973	2.5810	3.0641
16.0	1.8899	2.1201	2.6371	3.1663
17.0	1.9006	2.1403	2.6894	3.2653
18.0	1.9095	2.1579	2.7381	3.3614
19.0	1.9171	2.1737	2.7837	3.4548
20.0	1.9235	2.1870	2.8263	3.5457
∞	1.9597	2.2942	3.5714	∞

APPENDIX II

 SECOND MOMENTS ASSOCIATED WITH I_c AND I_s

The in-phase and quadrature components of the noise current I_N

$$\begin{aligned}
 I_c(t) &= \sum_{n=1}^M c_n \cos [(\omega_n - q)t - \varphi_n] \\
 I_s(t) &= \sum_{n=1}^M c_n \sin [(\omega_n - q)t - \varphi_n]
 \end{aligned}
 \tag{3.3}$$

are closely related to the envelope R and phase angle θ of the total current, this relationship being shown by the equations (3.4) and (3.5). $I_c(t)$ and $I_s(t)$ and their time derivatives may be regarded as random variables. In much of our work we have to deal with the probability distribution of these random variables. By virtue of the representation (3.3) and the central limit theorem¹⁸ this distribution is normal in the several variables. The coefficients in the quadratic form occurring in the exponent are deter-

¹⁸ Section 2.10 of Reference A.

mined by the second moments of the variables.¹⁹ Here we state these moments. Some of the moments have already been given in Sections 3.7 and 3.8 of Reference A. For the sake of completeness we shall also give them here. The new results given below are derived in much the same way as those given in Reference A.

Let

$$\begin{aligned}
 b_n &= (2\pi)^n \int_0^\infty w(f)(f - f_q)^n df \\
 b_0 &= \int_0^\infty w(f) df = \psi_0 \\
 g &= \int_0^\infty w(f) \cos 2\pi(f - f_q)\tau df \\
 h &= \int_0^\infty w(f) \sin 2\pi(f - f_q)\tau df
 \end{aligned} \tag{A2-1}$$

and let g', g'', h', h'' denote the first and second derivatives of g and h with respect to τ . For example,

$$g' = -2\pi \int_0^\infty w(f)(f - f_q) \sin 2\pi(f - f_q)\tau df$$

Incidentally, in many of our cases $w(f)$ is assumed to be symmetrical about f_q . This introduces considerable simplification because $b_1, b_3, b_5, \dots, h, h', h'',$ reduce to zero.

The following table gives values of b_n 's and g for two cases of frequent occurrence

	Ideal band pass filter centered on f_q	Normal law filter centered on $f_q, f_q \gg \sigma$
$w(f)$	w_0 for $f_a < f < f_b$ and zero elsewhere	$\frac{\psi_0}{\sigma\sqrt{2\pi}} e^{-(f-f_q)^2/2\sigma^2}$
b_0	$w_0(f_b - f_a)$	ψ_0
b_2	$\pi^2 w_0(f_b - f_a)^3/3$	$4\pi^2 \sigma^2 \psi_0$
b_4	$\pi^4 w_0(f_b - f_a)^5/5$	$48\pi^4 \sigma^4 \psi_0$
g	$(\pi\tau)^{-1} w_0 \sin \pi(f_b - f_a)\tau$	$\psi_0 e^{-2(\pi\sigma\tau)^2}$

If we write I_c, I'_c, I''_c for $I_c(t), I'_c(t), I''_c(t)$, where the primes denote differ-

¹⁹ Section 2.9 of Reference A.

entiation with respect to t , and do the same for $I_s(t)$ and its derivatives we have, from Section 3.8 of Reference A,

$$\begin{aligned}\overline{I_c^2} &= \overline{I_s^2} = b_0, & \overline{I_c I_s} &= 0 \\ \overline{I_c I_s'} &= -\overline{I_c' I_s} = b_1, & \overline{I_c I_c'} &= \overline{I_s I_s'} = 0 \\ \overline{I_c'^2} &= \overline{I_s'^2} = -\overline{I_c I_c''} = -\overline{I_s I_s''} = b_2, & \overline{I_c' I_s'} &= \overline{I_c I_s''} = \overline{I_s I_c''} = 0 \\ \overline{I_c I_s''} &= -\overline{I_s I_c''} = b_3, & \overline{I_c'' I_c'} &= \overline{I_s'' I_s'} = 0 \\ \overline{I_c''^2} &= \overline{I_s''^2} = b_4, & \overline{I_c'' I_s''} &= 0\end{aligned}\quad (\text{A2-2})$$

When we deal with moments in which the arguments of the two variables are separated by an interval τ as in (see the last of equations (3.7-11) of Reference A)

$$\overline{I_c(t) I_s(t + \tau)} = h,$$

it is convenient to denote the argument t by the subscript 1 and the argument $t + \tau$ by 2. Then our example becomes

$$\overline{I_{c1} I_{s2}} = h$$

We shall need the following moments of this type.

$$\begin{aligned}\overline{I_{c1} I_{c2}} &= \overline{I_{s1} I_{s2}} = g, & \overline{I_{c1} I_{s2}} &= -\overline{I_{c2} I_{s1}} = h \\ \overline{I_{c1} I_{c2}'} &= \overline{I_{s1} I_{s2}'} = -\overline{I_{c1}' I_{c2}} = -\overline{I_{s1}' I_{s2}} = g' \\ \overline{I_{c1} I_{s2}'} &= \overline{I_{c2} I_{s1}'} = -\overline{I_{c1}' I_{s2}} = -\overline{I_{c2}' I_{s1}} = h' \\ \overline{I_{c1}' I_{c2}'} &= \overline{I_{s1}' I_{s2}'} = -g'', & \overline{I_{c1}' I_{s2}'} &= -\overline{I_{c2}' I_{s1}'} = -h''\end{aligned}\quad (\text{A2-3})$$

It should be remembered that in these equations the primes on the I 's denote differentiation with respect to t while the primes on g and h denote differentiation with respect to τ .

APPENDIX III

EVALUATION OF A MULTIPLE INTEGRAL

Several multiple integrals encountered during the preparation of this paper were initially evaluated by the following procedure. The integral was first converted into a multiple series by expanding a portion of the integrand and integrating termwise. It was found possible to sum these series when one of the factorials in the denominator was represented as a contour integral. This reduced the multiple integral to a contour integral and sometimes the latter could be evaluated.

We shall illustrate this procedure by examining the integral

$$I = \int_{-\pi}^{\pi} d\theta \int_0^{\infty} dx x \exp \left[-x^2 + 2a \cos \theta + 2bx \sin \theta + c \sin^2 \theta \right] \quad (\text{A3-1})$$

Expanding that part of the exponential which contains the trigonometrical terms and integrating termwise gives

$$I = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{a^{2n} b^{2m} c^{\ell} \pi \Gamma(\ell + m + \frac{1}{2})}{n! \ell! (\ell + m + n)! \Gamma(m + \frac{1}{2})}$$

where we have used

$$2^{2n} \Gamma(n + \frac{1}{2}) n! = \sqrt{\pi} (2n)!$$

We next make the substitution

$$\frac{1}{(\ell + m + n)!} = \frac{1}{2\pi i} \int_C \frac{e^t dt}{t^{\ell+m+n+1}} \quad (\text{A3-2})$$

where the path of integration C is a circle chosen large enough to ensure the convergence of the series obtained when the order of summation and integration is changed. The summations may now be performed:

$$\begin{aligned} I &= \frac{1}{2i} \int_C dt e^{t+a^2/t} \sum_{m=0}^{\infty} b^{2m} t^{-m-1} (1 - ct^{-1})^{-m-1/2} \\ &= \frac{1}{2i} \int_C \frac{t^{-1/2} (t - c)^{1/2}}{t - c - b^2} e^{t+a^2/t} dt \end{aligned} \quad (\text{A3-3})$$

C encloses the pole at $c + b^2$ and the branch point at c as well as the origin.

When a^2 is zero the integral may be reduced still further. Let c be complex and b such that the point $c + b^2$ does not lie on the line joining 0 to c . Deform C until it consists of an isolated loop about $c + b^2$ and a loop about 0 and c , the latter consisting of small circles about 0 and c joined by two straight portions running along the line joining 0 to c . The contributions of the small circles about 0 and c vanish in the limit. Along the portion starting at 0 and running to c , $\arg(t - c) = -\pi + \arg c$, and along the portion starting at c and running to 0, $\arg(t - c) = \pi + \arg c$. On both portions $\arg t = \arg c$. Bearing this in mind and setting $t = c \sin^2 \theta$ on the two portions gives

$$I_{a=0} = \pi b (c + b^2)^{-1/2} e^{c+b^2} + 2c \int_0^{\pi/2} \frac{\cos^2 \theta e^{c \sin^2 \theta}}{b^2 + c \cos^2 \theta} d\theta \quad (\text{A3-4})$$

The integral may be expressed in terms of the function

$$Ie(k, x) = \int_0^x e^{-u} I_0(ku) du$$

by noting that

$$\begin{aligned} \int_0^1 \frac{e^{-\alpha - \beta \cos v}}{\alpha + \beta \cos v} dv &= \int_0^\pi dv \left[\frac{1}{\alpha + \beta \cos v} - \int_0^1 e^{-t(\alpha + \beta \cos v)} dt \right] \\ &= \pi(\alpha^2 - \beta^2)^{-1/2} - \pi \int_0^1 e^{-\alpha t} I_0(\beta t) dt \\ &= \pi(\alpha^2 - \beta^2)^{-1/2} - (\pi/\alpha) Ie(\beta/\alpha, \alpha) \end{aligned} \quad (\text{A3-5})$$

Thus

$$I_{a=0} = \pi e^{c/2} I_0(c/2) + (\pi b^2/\alpha) e^{b^2+c} Ie\left(\frac{c}{2\alpha}, \alpha\right) \quad (\text{A3-6})$$

where

$$\alpha = b^2 + c/2 \quad (\text{A3-7})$$

Noise in Resistances and Electron Streams

By J. R. PIERCE

TECHNICALLY correct results in a field are achieved initially in diverse and often confusing and complicated ways. Sometimes, such results are later brought together to give them a more unified form and a sounder basis; such critical summary and exposition is of great value. In quite another way, a worker who uses results established in a field will discover many plausible reasons for believing the results, and he will find eventually that an air of inevitability and "understanding" pervades the subject. Such "understanding" is not to be confused with the process of rigorous proof carried out step by step, but it can help in organizing and making use of a body of related material.

The field of "noise", especially as it affects electron devices and communications in general, is one particularly troublesome to engineers. The sound work on the subject has commonly involved mathematics and especially statistical ideas unfamiliar to many who must deal with the practical problems of noise. In early papers on noise, a great deal of heat was generated in acrimonious controversy between two schools, one of which assigned a uniform noise spectrum to certain noise sources, while the other held this to be inadmissible and got identical answers by more recondite means. Happily, a recent paper by S. O. Rice¹ clearly presents both approaches. Rice's paper further provides a fine broad summary of noise problems together with considerable original material. It does not extend far into the field of electronics.²

The reader who has sufficient time could achieve a profound "understanding" of the circuit aspects of noise by reading Rice's paper. The understanding would involve familiarity with much mathematics useful in itself. To many engineers, however, this might prove a lengthy and painful process.

The writer proposes to present here a series of plausible arguments for believing certain facts about noise. Both simple circuit considerations and "electronic" effects (as, space charge reduction of noise) are included. The arguments presented are not intended to be original and it is not claimed that they are rigorous; they do seem to be easily understood, and to help in remembering and in using some important practical material. Starting points of the arguments, or "postulates", have been chosen on the basis of familiarity, not simplicity. No effort is made to point out all of the hidden assumptions in the arguments, but a few important ones are indicated.

An initial warning should be made that quantum effects treated in Nyquist's original paper on Johnson noise, but afterwards much neglected, are entirely disregarded here.

I. JOHNSON NOISE³

In 1926, in an investigation of amplifiers with exceedingly high grid resistances, J. B. Johnson discovered that a resistance acts as a noise generator having an open-circuit voltage with a mean square value

$$\overline{v^2} = 4kTRB. \quad (1)$$

Here and subsequently, lower case letters v and i will be used in referring to noise voltages and currents. In (1), $\overline{v^2}$ is the mean square value of noise voltage components of frequency lying in a small bandwidth B (sometimes,

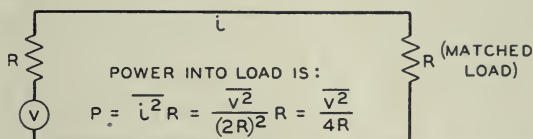


Fig. 1—Relations between noise power, noise voltage and noise current can be derived by assuming the noise source to be a voltage in series with a resistance.

called df or Δf), k is Boltzman's constant, and R is resistance. We easily see from Fig. 1 that the maximum noise power which can be made to flow from the resistance into a load (that which will flow into a matched load) is

$$P = \frac{\overline{v^2}}{4R} = kTB. \quad (2)$$

This "available noise power" is a convenient alternative formulation.

If an impedance has a reactive as well as a resistive component, the open circuit noise is given by (1) where R is the resistive component; if an admittance has a conductance G the noise may be represented as an impressed current (that which flows when the admittance is short circuited) of magnitude

$$\overline{i^2} = 4kTGB. \quad (3)$$

We see from (1) that if two resistances are connected in series, the total *squared* noise voltage is the sum of the *squares* of the noise voltages produced by the resistances separately, and from (3) we see that the noise currents of conductances connected in shunt also add by summing squares. This rule of addition holds for adding the noise of all independent sources. Of course, if noise from the same noise source reaches a point by different paths, the

voltage or current components near any frequency should be added directly with due regard for phase.

Johnson noise is related to many physically similar phenomena such as Brownian motion and the random fluctuations in position observed in the coils of very sensitive galvanometers.

The simplest derivation of (1), (2) or (3) is that given by Nyquist⁴ in a companion paper to Johnson's. Consider a long lossless transmission line of length L terminated at each end in resistances equal to its characteristic impedance. Imagine line and terminations in thermal equilibrium at a temperature T , as shown in Fig. 2. If electrical energy flows from the resistance at 1 to that at 2, then equal energy must then flow from 2 to 1, as any net gain or loss of energy would violate the second law of thermodynamics.

Now, suppose that we suddenly close the switches at 1 and 2, short circuiting the ends of the line. The line now becomes a resonator, having resonant

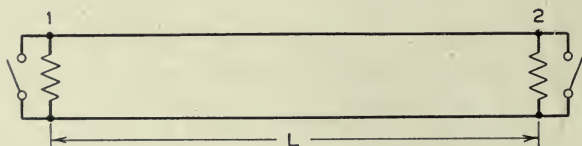


Fig. 2—Two resistances terminating a transmission line act as generators of thermal noise power traveling along the line.

frequencies such that the line is n half wavelengths long. The resonant frequencies will be

$$f = n(c/2L). \quad (4)$$

Here n is an integer and c is the velocity of light. The frequencies are separated by frequency intervals

$$\Delta f = (c/2L). \quad (5)$$

The energy which originally flowed right to left and left to right between the resistances is now reflected at the ends. It may be expressed as the thermal energy associated with the resonant modes of the line. According to statistical mechanics, there is an energy kT associated with each resonant mode. The energy per unit bandwidth is obtained by dividing this by the frequency interval between modes, given by (5) and is

$$w = kT/\Delta f = kT/(c/2L). \quad (6)$$

Since it takes a wave a time L/c to pass completely through the line, this energy w represents the energy per unit bandwidth which flowed into the

line from both resistances over a period L/c . If p is the power per unit bandwidth from one resistance, then

$$\begin{aligned} 2p(L/c) &= w = kT/(c/2L) \\ p &= kT. \end{aligned} \quad (7)$$

Or, we may say that the power flow from a resistance into a matched load (the available power) is, for a bandwidth B

$$P = kTB. \quad (8)$$

Sometimes it may be desired to know the mean squared fluctuation voltage integrated over all frequencies. Carrying out such an integration for the voltage between a pair of terminals connected by a complicated network would seem to be a difficult procedure. However, if the pair of terminals is shunted by a capacitance, the integrated fluctuation voltage can be obtained by direct application of the principles of statistical mechanics.

In a lumped network composed of capacitive, inductive and resistive elements* each capacitance and each inductance constitutes a degree of freedom; that is, the electrical state of the network can be specified completely by specifying the voltage across each capacitance and the current in each inductance**. According to statistical mechanics, the average stored energy per degree of freedom is $kT/2$. The stored energy in a capacitance is $Cv^2/2$. Thus, the mean squared noise voltage of all frequencies across a capacitance C must be

$$\overline{v^2} = kT/C. \quad (9)$$

Similarly, the mean squared noise current of all frequencies flowing in an inductance L is

$$\overline{i^2} = kT/L. \quad (10)$$

We have conveniently thought of Johnson noise as generated in the resistances in a network. We need not change this concept and say that the voltage and current of (9) and (10) are generated in the capacitance or inductance any more than we would say that the thermal velocities of molecules are generated by the molecules' mass. Relations (9) and (10) merely represent necessary consequences of the laws of statistical mechanics as, indeed, does (1).

It is of some interest to illustrate the use of (9) and its connection with (1)

* Strictly, such a lumped network is an unrealizable ideal. There are no pure capacitances, inductances, or resistances. The conditions under which actual condensers, coils and resistors can be represented satisfactorily by these idealizations must be judged by measurement or calculation or by past experience or intuition.

** In enumerating the degrees of freedom, capacitances in series or shunt are lumped together as one element; the same holds true for inductances.

by a very simple example. Consider a resonant circuit consisting of a capacitance C , an inductance L and a resistance R_0 , all in parallel. The resistive component of the impedance across this circuit is

$$R = \frac{R_0}{1 + Q^2 \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)^2} \quad (11)$$

$$Q = R_0 \omega_0 C = R_0 / \omega_0 L \quad (12)$$

$$\omega_0 = 1/\sqrt{LC}. \quad (13)$$

Here ω_0 is the resonant frequency of the circuit and Q has its usual meaning.

From (1) we see that as R_0 , the resistance at resonance ($\omega = \omega_0$) is made higher, the noise voltage for frequencies near resonance increases. However, if we regard ω_0 and C in (12) as fixed, we see that as R_0 is increased the Q of the circuit is increased, the frequency range over which R is high is decreased, and R actually becomes lower far from resonance. (9) tells us that the mean square noise voltage integrated over all frequencies remains constant as R_0 is changed.

It is found that for a high Q circuit, the noise is much like a carrier of frequency ω_0 modulated by low-frequency noise. If we let the radian frequency of this "noise modulation" be $(\omega - \omega_0)$, then the mean square amplitude of the noise modulation varies with frequency about as R given by (11) varies with $(\omega - \omega_0)$.

II. SCHOTTKY NOISE OR SHOT NOISE

In 1918 Schottky⁵ described the "Schrot-Effekt": the noise in vacuum tubes due to the corpuscular nature of the electron convection current. This is commonly known as "shot noise." The magnitude of this noise is usually derived by means quite different from those used here.

Johnson noise is necessarily associated with any electrical resistance, whatever its nature. Now, consider a close spaced planar diode shown in Fig. 3 consisting of two opposed emitting cathodes, each emitting a current I_0 . Suppose the whole diode is held at the same temperature. There are no batteries or other sources of power aside from thermal energy; the only electrical energy flow must then be Johnson noise, ascribable to the resistance of the diode.

Assume that the cathodes both have the same uniform work function. Then when the diode is short circuited, each electron emitted from cathode 1 will reach cathode 2, and each electron emitted from cathode 2 will reach cathode 1.* If cathode 2 were made negative, all the electrons from 2 would

* It is here assumed that I_0 is small enough so that depression of potential due to space charge is avoided.

continue to reach 1, but some of the low-velocity electrons leaving 1 would be turned back from 2.

It is well known⁶ that if a Maxwellian velocity distribution is assumed for the electrons leaving 1, the electrons which can overcome the retarding field and reach 2 are found to constitute a current

$$I = I_0 e^{eV/kT}. \quad (14)$$

Here I_0 is the total current carried by electrons leaving 1 and V is the voltage of 2 with respect to 1, which has been assumed to be negative.

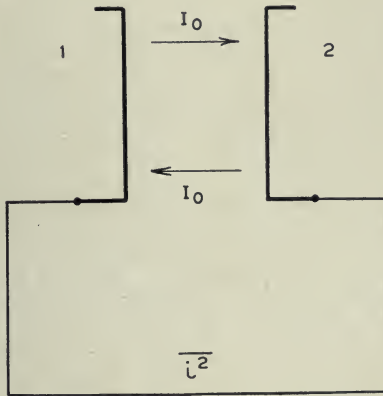


Fig. 3—An electronic resistance formed by two opposed cathodes at the same temperature acts as a generator of thermal noise.

By differentiating (14) we can obtain the diode conductance G at $V = 0$, and we find

$$G = \frac{e}{kT} I_0. \quad (15)$$

From (3) when the diode is short circuited and the voltage is zero we have a mean square noise current

$$\bar{i}^2 = 4kTGB = \frac{eI_0}{kT} (4kTB) = 4eI_0 B. \quad (16)$$

This noise is the sum of the noise due to two independent noise sources (the noise in the two currents I_0). That due to either current I_0 is*

$$\bar{i}^2 = 2eI_0 B. \quad (17)$$

* In this section, we are concerned with short transit angles only and no distinction need be made between the current induced in the circuit, i , and the electron convection current.

This is the expression for shot noise in a randomly emitted current, as in temperature limited emission or in photo electric emission.

III. NOISE OTHER THAN SHOT NOISE: ELECTRON MULTIPLIERS

Let us consider a class of systems in which the average output current is proportional to the average input current, in which an electron of charge, e entering produces an output charge, ne instantaneously, and in which the probability that any electron will produce n electrons is p_n .

If the input current is I_0 , the average output current is

$$I = \bar{n}I_0 \quad (18)$$

$$\bar{n} = \sum_n np_n. \quad (19)$$

It is easy to persuade ourselves that any frequency component of current, noise or signal, will produce an output current $\bar{n}$ times as great; this happens to be true, and we will use the fact.

Let us consider our device when it has randomly emitted electrons as an input. At the output we will see appear groups of 1, 2, 3 etc. electrons, each group caused by the entrance of a single electron. If I_0 is the total input current, the output current consisting of groups of n electrons is

$$I_n = nI_0p_n. \quad (20)$$

Each group carries a charge ne . We may now use (17) to write the noise in the part of the current carried by groups of n electrons, replacing the electronic charge, e , by the group charge, ne

$$\bar{i}_n^2 = 2(ne)(nI_0p_n)B. \quad (21)$$

As there is no correlation between entering electrons, the total mean square output noise current will be the sum of the noise components carried by groups consisting of different numbers n of electrons. Summing (19) with respect to n we obtain

$$\bar{i}_i^2 = 2eI_0B\bar{n}^2 \quad (22)$$

$$\bar{n}^2 = \sum_n n^2 p_n. \quad (23)$$

Now, the input has been taken as having shot noise. A part of the noise output is to be attributed to this input shot noise amplified by the device; that is, it will be $\bar{n}^2$ times the input shot noise.

$$\bar{i}_s^2 = \bar{n}^2 2eI_0 B. \quad (24)$$

The part of the noise output current due to the fact that an electron does

not produce $\bar{n}$ electrons, but may produce 0, 1, 2... etc. electrons, must be the difference between (22) and (24), or

$$\bar{i}_1^2 = 2eI_0 B(\bar{n}^2 - \bar{n}^2). \quad (25)$$

The quantity in parentheses is the mean square deviation in n .*

If the input current has any noise components $\bar{i}_n^2$, then the total noise output component will be

$$\bar{i}^2 = \bar{i}_1^2 + \bar{n}^2 \bar{i}_n^2. \quad (26)$$

By applying (26) successively to stage after stage the noise output of a multistage electron multiplier can be evaluated (if one knows $(\bar{n}^2 - \bar{n}^2)$).⁷

Wonder is sometimes expressed that current can be noisier than shot noise, in which the time of electron arrival is purely random. Obviously, we can have more than shot noise only if there is something non-random about the time of electron arrival, and the argument above discloses just what this is; it is the arrival of electrons in bunches.** We can easily see how erratic even large currents would be if electrons were bound together in groups having a total group charge of a coulomb, all the electrons in a group arriving simultaneously. Reverting to our shot noise formulas, we may illustrate this by assuming a perfect multiplier with a shot noise input, in which each input electron produces exactly N output electrons. Arguing from the shot noise equation (17) and replacing e by Ne we should expect an output noise current

$$\bar{i}^2 = 2(Ne)I_1 B \quad (27)$$

where I_1 is the output current; we get exactly the same result by assuming the input noise current squared amplified by N^2

$$\begin{aligned} \bar{i}^2 &= (2eI_0 B)N^2 \\ &= 2(Ne)(NI_0)B \\ &= 2(Ne)I_1 B. \end{aligned} \quad (28)$$

* The mean square deviation is the sum with respect to n of the square of the deviation from the mean value of n , $\bar{n}$.

$$\Sigma (n - \bar{n})^2 p_n = \Sigma n^2 p_n - 2\bar{n} \Sigma n p_n + \bar{n}^2 \Sigma p_n.$$

The summation in the first term is $\bar{n}^2$, that in the second term is $\bar{n}$ and that in the third term is unity. Hence

$$\Sigma (n - \bar{n})^2 p_n = (\bar{n}^2 - \bar{n}^2).$$

** Anything, (such as transit time difference for electrons within a bunch) which tends to break up the bunches will reduce the noise—and the signal as well. Such noise reduction involves a return to a more nearly random flow.

Conversely, we are led to wonder whether a current less bunched than that produced by random emission might not have less noise. The most smoothly distributed current we can imagine is that of f_0 electrons per second emitted at evenly spaced intervals. Obviously, such a current will have a spectrum consisting of frequencies nf_0 , integral multiples of f_0 . Thus for $f < f_0$, there will be no "noise" and similarly for $f_0 < f < 2f_0$, $2f_0 < f < 3f_0$, etc.

For a current of 10 *ma*, $f_0 = 6.3 \times 10^{16}$; thus, even for small currents an evenly spaced emission would have no a-c components in the radio-frequency range; this is a comforting thought in considering space-charge reduction of noise, which is discussed in section 5. However, purely to satisfy our curiosity we may pursue the matter a little further. If we assume that each electron constitutes an instantaneous pulse of current, a simple harmonic analysis shows that the a-c current component of frequency nf_0 will have a mean square value

$$\overline{i_n^2} = 2eI_0f_0. \quad (29)$$

Thus, in each interval f_0 wide centered about a frequency nf_0 there will be a mean squared a-c current equal to that which would be associated with the same band for random emission with the same current. By making the emission regular we have not reduced the mean square "noise" current in a broad frequency range; we have merely changed its frequency distribution from a uniform distribution to a distribution of sharp, high peaks.

IV. PARTITION NOISE

Consider a tetrode, shown in Fig. 4, with a cathode current I_c , a screen current I_s , and a plate current I_p .

The grid current is taken as zero. Suppose that the screen is very fine, so that every electron leaving the cathode has the same chance of striking the screen, regardless of its point of departure. We may now regard the function of the screen as that of a peculiarly simple electron multiplier, for which n can be zero (electron striking screen) or 1 (electron passing screen).

The probability of an electron passing the screen is I_p/I_c . Accordingly, from (19) and (23),

$$\bar{n} = I_p/I_c \quad (30)$$

$$\overline{n^2} = I_p/I_c. \quad (31)$$

Suppose we write the noise in the cathode current as

$$\overline{i^2} = \Gamma^2 2eI_c B \quad (32)$$

Here Γ^2 , a factor less than unity, is introduced to account for the "space charge noise reduction" in space charge limited flow.

Now, by applying (25) and (26) we obtain for the noise in the plate current

$$\begin{aligned}\overline{i_p^2} &= 2eI_c B(I_p/I_c - (I_p/I_c)^2) + \Gamma^2 2eI_c B(I_p/I_c)^2 \\ \overline{i_p^2} &= 2eI_p B(1 - (1 - \Gamma^2)(I_p/I_c)).\end{aligned}\quad (33)$$

It is to be noted that if $\Gamma^2 = 1$, that is, if the cathode current is random, the noise in the plate current is purely shot noise. The screen cannot make the plate current noisier than shot noise since it does not act to produce bunches of electrons.

The noise in the screen current can be obtained by substituting I_s for I_p

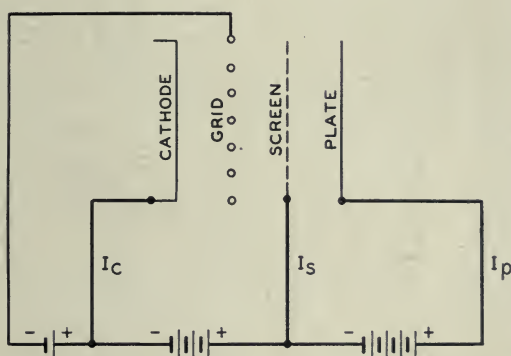


Fig. 4—Electrons randomly hitting or missing the screen grid make a tetrode noisier than a triode.

in (33). There is a correlation between the screen and plate noise currents; the total noise in the screen current plus the plate current must, of course, be

$$\overline{i_s^2} + \overline{i_p^2} = \overline{i_c^2} = \Gamma^2 2eI_0 B \quad (34)$$

and not the sum of $\overline{i_p^2}$ and $\overline{i_s^2}$.

Partition noise has been discussed by Thompson, North and Harris.⁸

V. SPACE CHARGE REDUCTION OF NOISE

In this section an approximate derivation of noise in a space charge limited diode will be presented. The derivation leads to an expression valid for many practical tubes and illustrates the nature of the noise in space charge limited flow.

Consider a parallel plane diode of unit area and spacing x , with an applied voltage V_0 , as shown in Fig. 5. When the voltage is applied, the electron convection current in the diode rises to value I_0 . Neglecting thermal velocities of electron emission, this current is such that the electronic

"space charge" associated with it causes the voltage gradient at the cathode surface to be zero. A greater current would mean a negative gradient at the cathode and hence no emission; a smaller current would mean a positive gradient at the cathode and unlimited emission. On this basis Child's law is derived, which gives the current per unit area I_0 in amperes in terms of the voltage V_0 and the spacing in centimeters x as

$$I_0 = (2.33) 10^{-6} V_0^{3/2} / x^2. \quad (35)$$

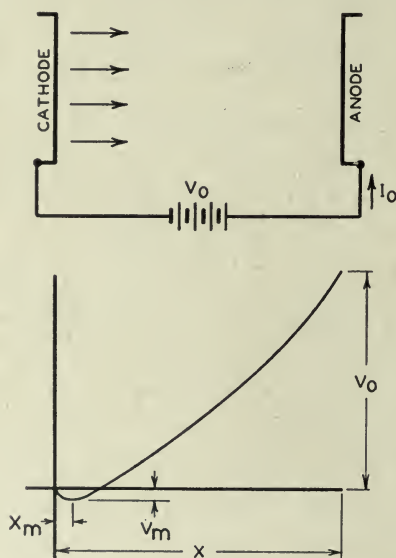


Fig. 5—Part of the electrons leaving the cathode of a diode are turned back before reaching the potential minimum; others proceed to the anode. Ordinarily the greater amount of noise is associated with the space between the potential minimum and the anode.

From (35) we can obtain a useful relation for the conductance G

$$G = \partial I_0 / \partial V_0 = (3/2)(I_0 / V_0). \quad (36)$$

The resistance R is

$$R = 1/G = (2/3)(V_0 / I_0). \quad (37)$$

In actual diodes, the electrons are emitted from the cathode with a thermal velocity distribution; a potential minimum of some negative voltage V_m is formed at some distance x_m from the cathode surface. If the magnitude of the emitted electron current is I_e and the actual current passing the potential minimum is I_0 , then because of the Maxwellian velocity distribution we have

$$\begin{aligned} I_0 &= I_e e^{eV_m/kT} \\ &= I_e e^{11,600 V_m/T}. \end{aligned} \quad (38)$$

Ordinarily, the magnitude of V_m is very small compared with V_0 ; I_0 is very small compared with I_c and x_m is very small compared with x .

Suppose V_m were held constant, say, by putting a conducting plane of potential V_m at x_m . Then, the electrons which pass this plane are quite independent of the low energy electrons which are turned back, and hence in the current passing x_m there will be pure shot noise.

$$\overline{i^2} = 2eI_0B. \quad (39)$$

Now suppose we change V_m . The change in I_0 will be, from (38),

$$dI_0 = dV_m/R_m \quad (40)$$

$$R_m = (eI_0/kT)^{-1}. \quad (41)$$

If we use a constant current instead of a constant voltage d-c supply, then V_m must fluctuate in such a way as to cause a current equal and opposite to (39), or, there must be a fluctuating voltage $\overline{v_m^2}$ such that

$$\begin{aligned} \overline{v_m^2} &= 2eI_0BR_m^2 \\ &= (1/2)4kTR_mB. \end{aligned} \quad (42)$$

Suppose we consider the noise fluctuation of the anode voltage of a space charge limited diode supplied from a constant-current source. If there were no fluctuations in the voltage drop between the potential minimum at x_m and the anode at x , (42) would give the noise voltage fluctuation of such an "open circuited" diode. Actually, much larger fluctuation voltages are observed, and we must conclude that they arise in the space between the potential minimum and the anode. As the current is constant in this region (by definition—we have assumed a constant-current supply) we are forced to conclude that such fluctuations are due to a variation of mean electron speed in this region. The field at x_m is necessarily zero. If, with a constant current, electrons travel more rapidly between x_m and the anode, there is less electronic charge everywhere in this region, the rate of change of field with distance, and hence, the field, are everywhere smaller, and the voltage between x_m and the anode at x will be smaller.

It is somewhat involved to treat the problem of multi-velocity flow exactly; this has been done by Rack⁹ and others^{8,10,11}; however, Rack has shown that an approximate treatment yields very nearly the correct result over a fairly wide range of conditions. In this approximation, the stream of electrons with many velocities and a fluctuating mean velocity is replaced by a stream in which all electrons have the same velocity, and this has a mean square fluctuation equal to that of the multi-velocity stream.

Let us now measure x from the potential minimum. Suppose we consider an electron which passed the potential minimum ($x = 0$) at $t = 0$.

The field at the potential minimum is zero. The charge which has flowed in behind the electron at the time t is $-t I_0$. Hence, from Gauss's theorem the potential gradient is

$$\partial V / \partial x = I_0 t / \epsilon \quad (43)$$

where ϵ is the dielectric constant of vacuum. We have for the acceleration

$$\ddot{x} = \frac{e}{m} \frac{I_0 t}{\epsilon} \quad (44)$$

If at the time $t = 0$ (at the potential minimum), $x = 0$, $\dot{x} = \dot{x}_0$

$$\dot{x} = \frac{e}{m} \frac{I_0}{2\epsilon} t^2 + \dot{x}_0 \quad (45)$$

$$x = \frac{e}{m} \frac{I_0}{6\epsilon} t^3 + \dot{x}_0 t \quad (46)$$

Now the voltage V between the potential minimum and any point x must be such that

$$\dot{x}^2 - \dot{x}_0^2 = 2 \frac{e}{m} V_0 \quad (47)$$

$$V_0 = \frac{1}{2 \frac{e}{m}} \left(\frac{e}{m} \frac{I_0}{2\epsilon} t^2 + \dot{x}_0 \right)^2 - \frac{I_0}{2\epsilon} t^2 \dot{x}_0 \quad (48)$$

At any fixed point x , if we vary $\dot{x}_0$ by a small amount $d\dot{x}_0$, we find by differentiating (46)

$$\frac{dt}{d\dot{x}_0} = - \frac{t}{\left(\frac{e}{m} \frac{I_0}{2\epsilon} t^2 + \dot{x}_0 \right)} \quad (49)$$

From (48)

$$dV_0 = \frac{I_0 t}{\epsilon} \left(\frac{e}{m} \frac{I_0}{2\epsilon} t^2 + \dot{x}_0 \right) dt + \frac{I_0}{2\epsilon} t^2 d\dot{x}_0 \quad (50)$$

Using (49)

$$dV_0 = - \frac{I_0}{2\epsilon} t^2 d\dot{x}_0 \quad (51)$$

It now remains to evaluate t . For most cases, the thermal velocities at the potential minimum are so small compared with the velocities in most of the region between the minimum and the anode that we can take the value of t for $\dot{x}_0 = 0$. Then, from (45) and (47)

$$t^2 = \left(\frac{e}{m} \frac{I_0}{2\epsilon} \right)^{-1} \left(2 \frac{e}{m} V_0 \right)^{1/2} \quad (52)$$

From (51) and (52)

$$dV_0 = -2^{1/2} \left(\frac{e}{m} \right)^{-1/2} V_0^{1/2} d\dot{x}_0. \quad (53)$$

Now, if $\overline{(d\dot{x}_0)^2}$ is the mean square fluctuation in velocity, the mean square fluctuation in voltage will be

$$\overline{v^2} = 2(e/m)^{-1} V_0 \overline{(d\dot{x}_0)^2}. \quad (54)$$

The assumptions leading to (54) are those leading to Child's law, and thus we can use (37) in connection with (54), giving

$$\overline{v^2} = 3(e/m)^{-1} I_0 R \overline{(d\dot{x}_0)^2}. \quad (55)$$

It now remains to evaluate $\overline{(d\dot{x}_0)^2}$, the mean square fluctuation in the velocity of the electrons passing the potential minimum; to do this, we return to (25). Suppose N is the number of input electrons per second. The output current can then be written

$$I_1 = \bar{n}Ne \quad (56)$$

and we can call the fluctuation in it

$$\overline{i^2} = \overline{(\delta \bar{n} Ne)^2}. \quad (57)$$

Equation (25) applies for no fluctuation in I_0 and hence for no fluctuation in N ; e is a constant, and thus we may write (25) as

$$\overline{(\delta \bar{n})^2} = \frac{2B}{N} (\overline{n^2} - \bar{n}^2). \quad (58)$$

We may generalize this to say that each electron has a probability p of producing some effect of magnitude n and the fluctuation in the magnitude of the effect is $\overline{(\delta \bar{n})^2}$. Before, we said that an electron had a probability p of producing n secondaries. Now we will say instead that an electron has an uncorrelated probability p of having a velocity u , and obtain for the mean fluctuation in the velocity, $\overline{(d\dot{x}_0)^2}$

$$\overline{(d\dot{x}_0)^2} = \frac{2B}{N} (\overline{u^2} - \bar{u}^2). \quad (59)$$

In a Maxwellian distribution, the number of electrons passing a plane perpendicular to the direction of motion per second having velocities lying in the range du at u is

$$dn = Aue^{-(mu^2/2kT_e)} du. \quad (60)$$

Here T_c is cathode temperature. We see $\bar{u}$ and $\bar{u^2}$ are

$$\bar{u} = \frac{\int_0^\infty u^2 \epsilon^{-(mu^2/2kT_c)} du}{\int_0^\infty u \epsilon^{-(mu^2/2kT_c)} du} = \sqrt{\frac{\pi}{2}} \sqrt{\frac{kT_c}{m}} \quad (61)$$

$$\bar{u^2} = \frac{\int_0^\infty u^3 \epsilon^{-(mu^2/2kT_c)} du}{\int_0^\infty u \epsilon^{-(mu^2/2kT_c)} du} = 2 \frac{kT_c}{m} \quad (62)$$

Accordingly

$$\bar{u^2} - \bar{u}^2 = \frac{1}{2}(4 - \pi) \frac{kT_c}{m} \quad (63)$$

Combining (63) with (59) we obtain

$$(\overline{d\dot{x}_0})^2 = \frac{B}{N} (4 - \pi) \frac{kT_c}{m} \quad (64)$$

Combining (62) with (53) and remembering that $I_0 = Ne$ we find the mean square open circuit noise voltage to be

$$\begin{aligned} \bar{v^2} &= 3(4 - \pi) kT_c R B \\ &= (.644) 4kT_c R B. \end{aligned} \quad (65)$$

This is the chief contribution to noise in a space charge limited diode.

Usually R is substantially equal to the plate resistance of the diode (it does not include effects on the cathode side of the potential minimum). Hereafter R will be treated as the total plate resistance of the diode.

VI. NOISE IN TRIODES AND PENTODES

Consider the triode shown in Fig. 6. Here we have a cathode, a grid, and a plate. The input admittance of the tube is represented in the diagram by the grid-cathode capacitance C_1 and the grid-plate capacitance C_2 . The resistance R_n is a fictitious noise resistance which will be evaluated later. It is assumed to act between the input admittance of the tube and the controlling action of the grid; no current can flow in R_n because the grid as indicated in the diagram is presumed to present an open circuit.

We will regard the cathode-grid region of the triode as an "equivalent diode" The anode voltage of the diode is taken as

$$V_0 = (V_g + V_p/\mu). \quad (66)$$

Here V_g is the grid voltage and V_p the plate voltage of the triode. If the plate voltage is held constant and μ is taken as constant

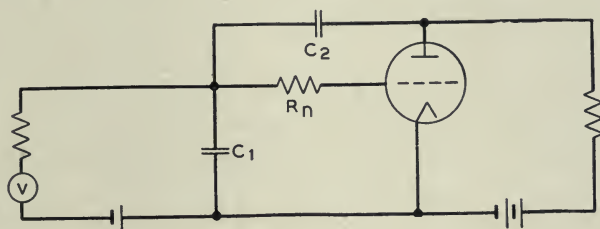
$$dV_0 = dV_g. \quad (67)$$

Hence, under these conditions

$$\partial I_0 / \partial V_0 = G = \partial I_0 / \partial V_g. \quad (68)$$

Here G is the conductance of the equivalent diode, the reciprocal of R which appears in (65), and is also the transconductance of the triode.

As we wish to calculate the noise with no a-c grid or plate voltage, and as these through (64) specify the plate voltage V_0 of the equivalent diode,



C_1 GRID-CATHODE CAPACITANCE
 C_2 GRID-PLATE CAPACITANCE

Fig. 6—Low-frequency noise in a triode can be ascribed to a fictitious noise resistance R_n , acting into an open circuit to cause voltage fluctuations on the grid.

the equivalent diode may be regarded as short-circuited. Hence, the noise current will be

$$\begin{aligned} \bar{i}^2 &= \bar{v}^2 / R^2 \\ &= (.644) 4kT_a GB. \end{aligned} \quad (69)$$

If we express this as shot noise reduced by a factor Γ^2 we obtain

$$\begin{aligned} \bar{i}^2 &= 2eI_0 \Gamma^2 B \\ \Gamma^2 &= (.644) \frac{2kT_c G}{eI_0}. \end{aligned} \quad (70)$$

Often, the noise expressed by (69) is ascribed as a fictitious noise resistance R_n , at room Temperature T , connected between the grid-cathode capacitance and the "controlling action" of the grid as shown in Fig. 6. This fictitious resistance looks into a complete open circuit; hence, it has a noise voltage

$$\bar{v}^2 = 4kTR_n B \quad (71)$$

and produces a noise plate current (for zero load resistance)

$$\bar{i}^2 = 4kTR_nBG^2. \quad (72)$$

Comparing (69) with (72) we find

$$R_n = (.644/G) (T_c/T_0). \quad (73)$$

Here T_0 is a reference temperature, usually taken as 290° K. The effect of load impedance on signal from this fictitious resistance is treated by purely circuit means.

In pentodes there is noise according to (69) and in addition there is partition noise according to (33). By taking Γ^2 from (70) and equating the noise current given by (33) to (72) the fictitious "noise resistance" of a pentode can be evaluated in terms of g , I_p/I_c and T_c/T .

REFERENCES

1. S. O. Rice, "Mathematical Analysis of Random Noise," *B. S. T. J.*, XXIII, pp. 282-332, July 1944, and XXIV, pp. 46-156, January, 1945.
2. D. A. Bell has given a valuable summary and discussion of noise and especially of noise in vacuum tubes. "Fluctuations of Electric Current," *Journal of the I. E. E.*, Vol. 93, pt. 3, pp. 37-44, January 1946.
3. J. B. Johnson, "Thermal Agitation of Electricity in Conductors," *Phys. Rev.*, Vol. 32, No. 1, pp. 97-109, July (1928).
4. H. Nyquist, "Thermal Agitation of Electric Charge in Conductors," *Phys. Rev.*, Vol. 32, No. 1, pp. 110-113, July (1928).
5. W. Schottky, "Spontaneous Current Fluctuations in Various Conductors," *Ann. Physik*, 57, pp. 541-567, (1918).
6. Irving Langmuir and Karl T. Compton, "Electrical Discharge in Gases," Part II, *Rev. Mod. Phys.*, Vol. 3, #2, p. 226 (April 1931).
7. W. Shockley and J. R. Pierce, "A Theory of Noise for Electron Multipliers," *Proc. I. R. E.* 26, pp. 321-332, March 1938.
8. B. J. Thompson, D. O. North and W. A. Harris, "Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies"
 Part I—*R. C. A. Review*, Jan. 1940
 Part II— " Apr. and July 1940
 Part III— " Oct. 1940
 Part IV— " Jan. 1941
 Part V— " Apr. and July 1941.
9. A. J. Rack, "Effect of Space Charge and Transit Time on the Shot Noise in Diodes," *B. S. T. J.*, Vol. 17, No. 4, p. 592, October (1938).
10. W. Schottky, and E. Spenke, "The Space-Charge Reduction of Shot Effect," *Wissenschaftliche Veröffentlichungen aus den Siemens-Werken*, Vol. 16, No. 2, pp. 1-41, July (1937).
11. E. Spenke, "The Space-Charge-Reduction of Shot Effect," *Wiss. Veröff. aus den Siemens-Werken*, Vol. 16, No. 2, p. 19, July (1937).

Abstracts of Technical Articles by Bell System Authors

*Gross-linkage of Linear Polyesters by Free Radicals.*¹ W. O. BAKER. Reactions fundamental to the use of the new low pressure laminating or casting resins have been studied. The striking property of these plastics, which are usually based on some polyester and a vinyl monomer, is their rapid and easy curing, leading to unique ease of fabrication. This curing, the formation of a permanent three-dimensional polymer network, or gel, is achieved by reaction with a source of free radicals, such as from an organic peroxide. These agents cause polymerization of the vinyl monomer, as was previously understood, but they also seem to incorporate the polyester into the network, even if the polyester contains little or no unsaturation.

Investigation of a series of simple polyesters, the polyundecanoates, showed that free radicals, such as come from the decomposition of benzoyl peroxide, could cross-link or gel the linear, saturated, chains. Apparently the hydrogen atoms in methylene groups next to polar groups like the carbonyl, i.e., the α -hydrogens, are removed by the free radicals. The resulting chain radical attacks an adjacent chain, and a cross-link is formed. The effects of cross-links thus produced on solubility, dilute solution viscosity, melt viscosity and, finally, stress relaxation of the cured solid were examined. Probably the similar activity of α -hydrogen atoms is important in the chemical aging or weathering of plastics and rubbers. It is likewise significant for the vulcanization of many synthetic rubbers.

*Rubberlike Products from Linear Polyesters.*² B. S. BIGGS, R. H. ERICKSON and C. S. FULLER. The polymers which result from the condensation of dibasic acids with propylene glycol are viscous gums which can be vulcanized to rubberlike products. In the unpigmented condition these rubbers are quite weak, but when reinforced with suitable pigments their strength and elongation compare favorably with other synthetic rubbers. Because polyesters of known structure and molecular weight can be easily synthesized, these polymers are useful for the study of the relations between structure and properties in rubberlike materials in general. Factors affecting tensile strength, oil resistance, brittle temperature, and stability are discussed.

*Pulse Code Modulation.*³ H. S. BLACK and J. O. EDSON. A radically new modulation technique for multichannel telephony has been developed which involves the conversion of speech waves into coded pulses. This new tech-

¹ *Jour. Amer. Chemical Soc.*, May 1947.

² *Indus. & Engg. Chem.*, September 1947.

³ *Telephony*, August 30, 1947.

nique is called Pulse Code Modulation or simply PCM. An eight-channel system embodying these principles was developed and produced in portable form for field operation. Other work carried on simultaneously by W. M. Goodall (see B. S. T. J., July 1947) resulted in the development of an experimental system using a different method of coding.

In carrying out this new type of modulation, the speech wave applied to each channel is, in effect, transmitted sample by sample, and each sample is represented by a multi-unit code employing on-or-off pulses, hence the term PCM.

This method appears to have exceptional possibilities from the standpoint of freedom from interference. Its full significance in connection with future radio and wire transmission may take some time to reveal.

*Stereoscopic Drawings of Crystal Structures.*⁴ W. L. BOND. A method is presented for getting stereoscopic pairs of atomic structure views given the coordinates of the atoms and cell constants.

*Properties of Liquids at High Sound Pressure.*⁵ H. B. BRIGGS, J. B. JOHNSON and W. P. MASON. When sound of high amplitude is transmitted into a liquid by means of a mechanical driving device, the ultimate limitation to the power that can be transferred is cavitation or breakdown of the liquid under high internal stresses. A study of cavitation has resulted in establishing the following results. Under steady-state conditions, light liquids filled with air cavitate when the negative acoustic pressure reaches the atmospheric pressure. When liquids are degassed, their natural cohesive pressure becomes effective and they will withstand a negative acoustic pressure. It is found that the total negative pressure required to cause cavitation is equal to the sum of the cohesive pressure—tensile strength—and the ambient pressure. Viscous liquids have a higher cohesive pressure and a proportionality has been established between the logarithm of the viscosity and the cohesive pressure. The amount of power that a liquid can withstand increases markedly as the pulse length is shortened.

An explanation of these phenomena is attempted on the basis of Eyring's theory of viscosity, plasticity and diffusion. On this theory natural holes exist in the liquid into which molecules can jump, leaving holes behind them. A jump occurs when the molecule has accumulated enough heat energy to surmount an activation potential barrier of energy value E_0 . Cavitation appears to be the result of coalescing of the natural holes in the negative pressure phase of the cycle. Since a molecule has to jump from a hole in order that this can coalesce with another hole, the cavitation pressure is proportional to the activation energy which in turn is proportional to the log-

⁴ *The American Mineralogist*, July-August 1947.

⁵ *Jour. Acous. Soc. Amer.*, July 1947.

arithm of the viscosity. The increased power-transmitting capacity for short pulse lengths is a result of the finite time taken for the small holes to grow in size to a large enough hole to cause rupture of the liquid.

*Modulation in Communication.*⁶ F. A. COWAN. The fundamentals involved in introducing signals into one medium and transmitting them through another are simplified in this review, so that the relationships between the many varieties of modulations attempted or in contemporary use are formed into a cohesive whole.

*Air-borne Magnetometers.*⁷ E. P. FELCH,* W. J. MEANS,* T. SLONCZEWSKI,* L. G. PARRATT, L. H. RUMBAUGH and A. J. TICKNER.* Developed under the impetus of the submarine menace of World War II, the air-borne magnetometer has found many peacetime uses. Navy airplanes equipped with magnetometers for exploration of Antarctica were used in the recent United States Navy expedition. An expedition now is studying the Aleutian Alaskan volcanos and the Aleutian submarine trench. From there it will proceed to Hawaii and Bikini.

*The Generation of Centimeter Waves.*⁸ H. D. HAGSTRUM. The electronic devices used most extensively, recently, for the generation of centimeter waves are discussed. The physical form, operating capabilities, and the basic physical principles of operation of the triode, velocity-variation, and magnetron oscillators are presented. An attempt is made to show how these oscillators are related to one another. For a variety of reasons, particular emphasis is placed on the magnetron oscillator.

*Selective Demodulation.*⁹ DONALD B. HARRIS. A method of demodulation is proposed in which the output current of the demodulator is a linear function of the input voltage, while at the same time provision is made for producing the necessary product terms which will result in demodulation. Demodulation is brought about by integrating the product of the instantaneous value of the modulated wave by the instantaneous value of a wave having the same frequency and phase as the carrier. Where this method of demodulation is used it is proposed that two carriers in quadrature on the same frequency may be employed, reducing the bandwidth to that required for single-sideband transmission.

It is suggested that the required linear demodulation characteristics may be obtained through the use of "electron-coupled" demodulators. Theoretical considerations indicate that, when demodulation of this type is employed, selectivity ahead of the demodulator may be dispensed with, the

⁶ *Elec. Engg.*, September 1947.

⁷ *Elec. Engg.*, July 1947.

* *Of the Bell System.*

⁸ *Proc. I. R. E.*, June 1947.

⁹ *Proc. I. R. E.*, June 1947.

signal-to-noise ratio is improved, greater economy of spectrum space is obtained, the number of tubes required is materially reduced through the use of a common intermediate-frequency amplifier for a number of channels, and any impairment due to the instability of the carrier or oscillator frequency is reduced.

As an example of the possible application of the principles outlined, a hypothetical eight-channel transmission system is described.

*The Physical Significance of Birkhoff's Gravitational Equations.*¹⁰ HERBERT E. IVES. Birkhoff's gravitational equations are put in terms of dt in place of the local time ds used by him. The transformed equations show that Lorentzian mass has been used, and to the Newtonian attractive force is added a force normal to the direction of motion, v^2/c^2 times the component of the gravitational force normal to the motion.

*Attenuation and Scattering of High-Frequency Sound Waves in Metals and Glasses.*¹¹ W. P. MASON and H. J. MCSKIMIN. By using a pulse method, attenuation and velocity measurements have been made for aluminum and glass rods in the frequency range from 2 to 15 megacycles. The sound pulses are generated by crystals waxed to the surface of the rod. This wax joint limits the band width of the transmitted pulse and measurements are made using long pulses which approach steady state conditions. The reflected pulses show evidence of several normal modes which can be minimized by using specially shaped electrodes. Longitudinal waves show delayed pulses of smaller magnitude that are caused by the longitudinal wave breaking up into reflected longitudinal and shear waves at the boundary. This effect is small if the diameter of the rod is 20 wave-lengths or more.

The measured losses for aluminum rods show a component proportional to the frequency and another component proportional to the fourth power of the frequency. The first component is the hysteresis loss found for most solid materials. The component proportional to the fourth power of the frequency is caused by Rayleigh scattering losses which are the result of differences in the elastic constants between adjacent grains caused by changes in orientation. Calculated scattering losses agree quite well with the measured values. The fourth-power scattering law holds quite well until the grain size is equal to one-third of a wave-length. For higher frequencies the scattering loss increases more nearly with the square of the frequency. Glasses and fused quartz have a loss directly proportional to the frequency, showing that any irregularities must be of very small size.

*The Growth of Auditory Sensation.*¹² W. A. MUNSON. The integration of sensation with respect to time was studied experimentally by means of tones

¹⁰ *Phys. Rev.*, August 1, 1947.

¹¹ *Jour. Acous. Soc. Amer.*, May 1947.

¹² *Jour. Acous. Soc. Amer.*, July 1947.

of short duration. Loudness tests were made on sounds persisting from 0.005 to 0.2 second and covering a wide range of levels. The observed increase in magnitude of a sensation as the duration time is increased is attributed to the integration characteristic of the central nervous system, and an equivalent electrical circuit is derived. The circuit analogy is then used in the computation of loudness as a function of the duration of the stimulus.

*The Physics of Electronic Semiconductors.*¹³ G. L. PEARSON. The band theory of solids is capable of explaining such fundamental properties of electronic semiconductors as the dependency of specific resistance on impurity content, the negative temperature coefficient of resistance, the sign of the Hall and thermoelectric effects, and the direction of rectification. Measurements of the specific resistance and the Hall constant enable the calculation of density, mobility, and mean free path of the electric carriers as a function of temperature and impurity.

*Automatic Frequency Control of Microwave Oscillators.*¹⁴ VINCENT C. RIDEOUT. A method for the automatic frequency control of any type of tunable microwave oscillator is described. In this method a servomechanism is used which includes a wave-guide discriminator circuit, a mercury-contact relay, a 60-cycle amplifier, and a small two-phase induction motor.

Tests made on a preliminary model of a circuit of this type used with a 4000-megacycle oscillator showed that a stability of one part in 50,000 was obtainable. The manner in which such a control system may be used in a microwave repeater is described.

*Proposed Method of Rating Microphones and Loudspeakers for Systems Use.*¹⁵ FRANK F. ROMANOW and MELVILLE S. HAWLEY. Proposed, is a method of rating microphones and loudspeakers whereby the over-all performance of a sound system may be determined by adding together the microphone and loudspeaker ratings and the gain of the interconnecting network. This sum gives the performance quite accurately for most systems. However, in some combinations of elements correction terms must be added. The formulas for these correction terms are derived.

The proposed microphone and loudspeaker ratings have the additional usefulness of being in a form which permits the comparison of instruments of different impedances.

*Sulfur Linkage in Vulcanized Rubber.*¹⁶ MILTON L. SELKER and A. R. KEMP. The reaction of 2-methyl-2-butene with sulfur at 141.6°C. was studied. Reaction time and concentration paralleled those common in

¹³ *Elec. Engg.*, July 1947.

¹⁴ *Proc. I. R. E.*, August 1947.

¹⁵ *Proc. I. R. E.*, *Waves and Electrons Section*, September 1947.

¹⁶ *Indus. & Engg. Chem.*, July 1947.

rubber-sulfur vulcanization. The results offer further insight into the vulcanization problem. The products of the reaction are liquids of the polysulfide type $R-S_x-R$, where x varies from 2 to 6 and R is an alkyl or alkenyl group and two solids ($C_5H_6S_3$ and a higher homolog). The polysulfides appear to be somewhat richer in hydrogen than is expected from reaction of two C_5H_{10} molecules, whereas the solids are hydrogen-poor. The structure of an acid anhydride in the sulfur system showing thione-thiol tautomerism is proposed for $C_5H_6S_3$, which is therefore 2,5-dithione-3-methyltetrahydrothiophene. The color changes with reaction time, from yellow to red to black, parallel those of rubber-sulfur vulcanizates. As in rubber-sulfur vulcanization the sulfur reaction rate is directly proportional to time, although the absolute rate is twice that in the polymer system. Starting with equal mole quantities of olefin and sulfur, there is a considerable amount of unreacted olefin in the system when all of the sulfur has reacted. The shorter the reaction time, the higher the value of x in the polysulfide $R-S_x-R$ and the larger the percentage of residues R that are saturated.

*On Hearing in Water vs. Hearing in Air.*¹⁷ L. J. SIVIAN. The paper deals with the ability of a submerged listener to hear sounds generated in the air above him, compared with their audibility when his head projects above the water. In a theoretical discussion it is shown that at 1000 c.p.s. a loss of the order of 45–55 db might be expected in the in-water audibility relative to the in-air value. This involves a number of assumptions, e.g., that there is no appreciable noise created by the listener's propulsion, and that the effect of hydrostatic pressure unbalance on the eardrum is negligible. A few measurements made at 1000 c.p.s. and 3000 c.p.s. yielded values which are not at variance with the theoretical analysis.

*Cathode Phase Inverter Design.*¹⁸ C. W. VADERSEN. Part I of this paper covers the general analysis of the cathode coupled phase inverter and develops formulae that enable the designer to compute circuit elements with good accuracy. The theory developed shows that degeneration exists only in the driven side of the amplifier and is limited to 6 decibels. Balance is discussed in terms of the tube parameters and external resistances, its being shown that considerable stability is attainable. A form of the inverter in which the power output stage utilizes a transformer with an unbalanced plate winding is presented. This is shown to give a true power balance in a manner analogous to the unbalanced plate resistor form of the voltage amplifying inverter.

Part II presents a graph of the general design equations and illustrates its use with several working examples.

¹⁷ *Jour. Acous. Soc. Amer.*, May 1947.

¹⁸ *Audio Engineering*, June and July 1947.

Contributors to this Issue

WINSTON E. KOCK, B.E., University of Cincinnati, 1932; M.S., 1933; Ph.D., University of Berlin, 1934. Institute for Advanced Study, Princeton, New Jersey, 1935-36. Director of Electronic Research, Baldwin Piano Company, Cincinnati, Ohio, 1936-42. Bell Telephone Laboratories, Research Department, 1942-. Dr. Kock was engaged in radar antenna work in the Radio Research Department during the war. He is now engaged in microwave and acoustic research.

W. D. LEWIS, A.B. in Communication Engineering, Harvard College, 1935; Rhodes Scholar, Wadham College, Oxford; B.A. in Mathematics, Oxford, 1938; Ph.D. in Physics, Harvard, 1941. Bell Telephone Laboratories, Inc., 1941-. Dr. Lewis was engaged in radar antenna work in the Radio Research Department during the war; he is now engaged in microwave repeater systems research.

L. A. MEACHAM, B.S. in Electrical Engineering, University of Washington, 1929; Certificate of Research, Cambridge University, England, 1930. Bell Telephone Laboratories, 1930-. From 1930 to 1941 Mr. Meacham's work dealt with crystal oscillators, multivibrators, phase shifters, and other devices used in precision standards of frequency. During the war he developed range measuring devices for radar, and has since been concerned with applications of pulse techniques to multiplex telephony.

E. PETERSON, Cornell University, 1911-14; Brooklyn Polytechnic, E.E. 1917; Columbia University, A.M. 1923; Ph.D. 1926. Electrical Testing Laboratories, 1915-17; Signal Corps, U. S. Army, 1917-19. Western Electric Company, Engineering Department, 1919-25; Bell Telephone Laboratories, 1925-. Lecturer in Electrical Engineering, Columbia, 1934-. As circuit research engineer, Dr. Peterson's work has been largely in theoretical studies of non-linear circuits and circuit elements.

J. R. PIERCE, B.S. in Electrical Engineering, California Institute of Technology, 1933; Ph.D., 1936. Bell Telephone Laboratories, 1936-. Engaged in study of vacuum tubes.

S. O. RICE, B.S. in Electrical Engineering, Oregon State College, 1929; California Institute of Technology, 1929-30, 1934-35. Bell Telephone

Laboratories, 1930-. Mr. Rice has been concerned with various theoretical investigations relating to telephone transmission theory.

V. C. RIDEOUT, B.Sc. in Engineering Physics, University of Alberta, 1938; M.S. in Electrical Engineering, California Institute of Technology, 1940. Bell Telephone Laboratories, Inc., 1939-1946. Department of Electrical Engineering, University of Wisconsin, 1946-. During the war Mr. Rideout worked in the Research Department on various components for radar systems; after the war he worked on frequency control systems and on intermediate frequency power amplifiers for microwave repeater systems.

R. W. SEARS, A.B. Ohio Wesleyan University, 1928; M.S., Ohio State University, 1929. Columbia University, 1930-1935. Bell Telephone Laboratories, 1929-. Mr. Sears has been engaged in research work on thermionics and semiconductors. Since 1939 he has been primarily concerned with the development of electron tubes.

L. C. TILLOTSON, B.S. in E.E., University of Idaho, 1938; M.S. in E.E., University of Missouri, 1941. Instructor in Electrical Engineering, University of Missouri, 1940-41. Bell Telephone Laboratories, Inc., 1941-. During the war Mr. Tillotson was engaged in the design and development of wave filters and other transmission networks. In 1946 he was transferred to the Radio Research Department and since that time has been concerned with microwave repeater systems research.

Public Library
Kansas City, Mo.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS
OF ELECTRICAL COMMUNICATION

Microwave Repeater Research.....*H. T. Friis* 183

The Measurement of Delay Distortion in Microwave
Repeaters.....*D. H. Ring* 247

Frequency Shift Telegraphy—Radio and Wire Applications
J. R. Davey and A. L. Matte 265

Reflections from Circular Bends in Rectangular Wave
Guides—Matrix Theory.....*S. O. Rice* 305

The Approximate Solution of Linear Differential Equations
Marion C. Gray and S. A. Schelkunoff 350

Potential Coefficients for Ground Return Circuits
W. Howard Wise 365

Abstracts of Technical Articles by Bell System Authors.. 372

Contributors to this Issue 377

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

THE BELL SYSTEM TECHNICAL JOURNAL

*Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.*



EDITORS

R. W. King

J. O. Perrine

EDITORIAL BOARD

W. H. Harrison

O. E. Buckley

O. B. Blackwell

M. J. Kelly

H. S. Osborne

A. B. Clark

J. J. Pilliod

F. J. Feely



SUBSCRIPTIONS

Subscriptions are accepted at \$1.50 per year. Single copies are 50 cents each.
The foreign postage is 35 cents per year or 9 cents per copy.



Copyright, 1948
American Telephone and Telegraph Company

The Bell System Technical Journal

Vol. XXVII

April, 1948

No. 2

Microwave Repeater Research

By H. T. FRIIS

INTRODUCTION

IT WAS some 80 years ago that Maxwell and Hertz demonstrated that free space is a good transmission medium for electromagnetic waves. Since this fundamental contribution, the radio art has advanced tremendously and a decade ago it had progressed to the point where it was possible to construct equipment suitable for quantitative propagation studies of microwaves. Such studies were made and they indicated that normal propagation over "line-of-sight" paths of signals of 10 to 20 centimeters wavelength was characterized by free space attenuation and freedom from atmospheric interference. These results, together with the facts that in this wavelength range wide bands of frequencies are available and it is possible to design small antennas having high directivity, encouraged us to start more comprehensive research work on microwave repeater circuits. This paper gives the present status of the work which was interrupted by our war efforts and resumed at the end of the war with the construction of an experimental New York-Boston system as an initial objective.

The first section will describe our propagation studies. It will be followed by sections on repeater circuit planning, antennas, radio frequency channel filters, the construction and testing of the repeater amplifier, and a concluding section on the whole repeater.

I. PROPAGATION STUDIES*

That portion of the radio frequency spectrum represented by wavelengths shorter than about five meters has long been considered as the proper domain for point-to-point communication links, local broadcasting, and mobile radio communication. Since these ultra-short waves are not reflected by the ionosphere, their effective range is not much greater than the horizon distance and it therefore becomes possible for a number of stations, properly separated, to operate in the same frequency band; for the same reason, atmospheric interference is not an important factor in this wave-

* This section was prepared by A. B. Crawford who, with W. M. Sharpless, is at present engaged in microwave propagation studies.

length range. Also, as the wavelength decreases, it becomes possible to construct antennas large in comparison with the wavelength so that high antenna gains are obtained and the corresponding directivity further reduces the interference areas.

Since about 1930, with the exception of the war years, we have conducted fundamental studies in radio propagation, taking advantage of advances in the art to extend the wavelength range from about four meters (ultra-short wave region) in the beginning to 1.25 centimeters (microwave region) at the present time. A considerable portion of the effort of those engaged in propagation studies has, of necessity, been devoted to the development of measurement techniques and reliable measuring apparatus. The present discussion, however, will be concerned with the results of experiments rather than with a description of the apparatus and methods. Most of these results have been described in the literature; the following is a review intended to show the development of the background leading to the present field trial of a microwave repeater circuit.

The object in making propagation studies has been to evaluate and to understand the effects of the terrain and of the lower atmosphere upon the transmission of ultra-short-wave and microwave signals. The evaluation is usually obtained by amassing sufficient data on a particular transmission experiment so that a statistical analysis can be made. Efforts to understand the transmission phenomena usually take the form of experiments involving specially designed apparatus. These experiments are varied from time to time as information is obtained or as it becomes desirable to check the validity of such theories as may be devised. The hope is always present that an understanding of the phenomena may suggest a means for reducing the transmission difficulties.

The absence of ionospheric reflections at these frequencies suggested at the start that propagation studies would probably be concerned mainly with phenomena familiar in optics, namely: reflection, refraction and diffraction. Two of the early papers^{1, 2} treated ultra-short-wave propagation from this viewpoint. It was soon observed that diffracted signals tended to be unstable in the shadow region; furthermore, as the wavelength is decreased the shadows cast by obstacles such as hills or the bulge of the earth itself become more sharply defined. For these reasons, a considerable part of our experimental work has been done on paths for which a line-of-sight exists between transmitter and receiver. The chief interest, therefore, has been in ground reflections and the effect of the atmosphere.

¹ J. C. Schelleng, C. R. Burrows and E. B. Ferrell, "Ultra-Short-Wave Propagation", *Proc. I. R. E.*, vol. 21, pp 427-463; March 1933.

² C. R. Englund, A. B. Crawford and W. W. Mumford, "Some Results of a Study of Ultra-Short Wave Transmission Phenomena", *Proc. I. R. E.*, vol. 21, pp 464-492; March 1933.

GROUND REFLECTIONS

Some of our first experiments with ultra-short-waves showed that regular reflections were obtained locally from open, relatively flat fields. The reflection coefficients were in good agreement with theory. Later, measurements of propagation between a transmitter located on a hill top and a receiver carried in an airplane² showed that for near-grazing angles of incidence, the irregular and wooded terrain, typical of the New Jersey countryside, could give rise to regular reflections at wavelengths as short as four meters. The depth of the minima in received signal strength, caused by wave interference between the direct and ground reflected components, corresponded to a reflection coefficient of about 0.9. In 1939, unpublished results obtained over the 39-mile Beer's Hill-Lebanon optical path (See map of Fig. I-1) indicated that for a wavelength of 30 centimeters the reflection coefficient was still large, about 0.8.

More recently, microwave propagation studies have been made over the same type of terrain at wavelengths of 3.25 centimeters and 1.25 centimeters and the situation in regard to ground reflections seems to have changed somewhat. Experiments were conducted over the 12.6 mile Beer's Hill-Deal path in which the height of the transmitting terminal was varied and which also made use of narrow-beam scanning antennas to separate the direct wave from a possible ground reflected component. The results showed the apparent reflection coefficient to be of the order of 0.2 at 3.25 centimeters and to be even less at 1.25 centimeters. Figure I-2 shows typical curves of signal level versus transmitter heights for wavelengths of 3.25 and 1.25 centimeters. Actually, the shapes of the curves can be accounted for better by diffraction, for which the hill about two miles from Deal is considered to be a straight edge, than by reflection from an assumed average ground plane. The true picture is probably a combination of reflection and diffraction effects.

In an effort to minimize ground reflection, over-water paths were avoided in the layout of the New York-Boston microwave repeater circuit and as a final check a number of variable antenna-height tests* were made in the preliminary survey of all sites. A few curves obtained at a wavelength of 7 centimeters are reproduced in Fig. I-3. Similar results were observed during a survey of sites between Chicago and Milwaukee.

It is concluded, therefore, that although in the wavelength range down to 30 centimeters, at least, the effects of ground reflection must be taken into account in the choice of sites for an optical path radio circuit, in the lower microwave range, below say 10 centimeters, scattering and absorption of the reflected wave by rough terrain and vegetation usually results in substantially free-space propagation under normal conditions when the line of

* F. F. Merriam was in charge of this work.

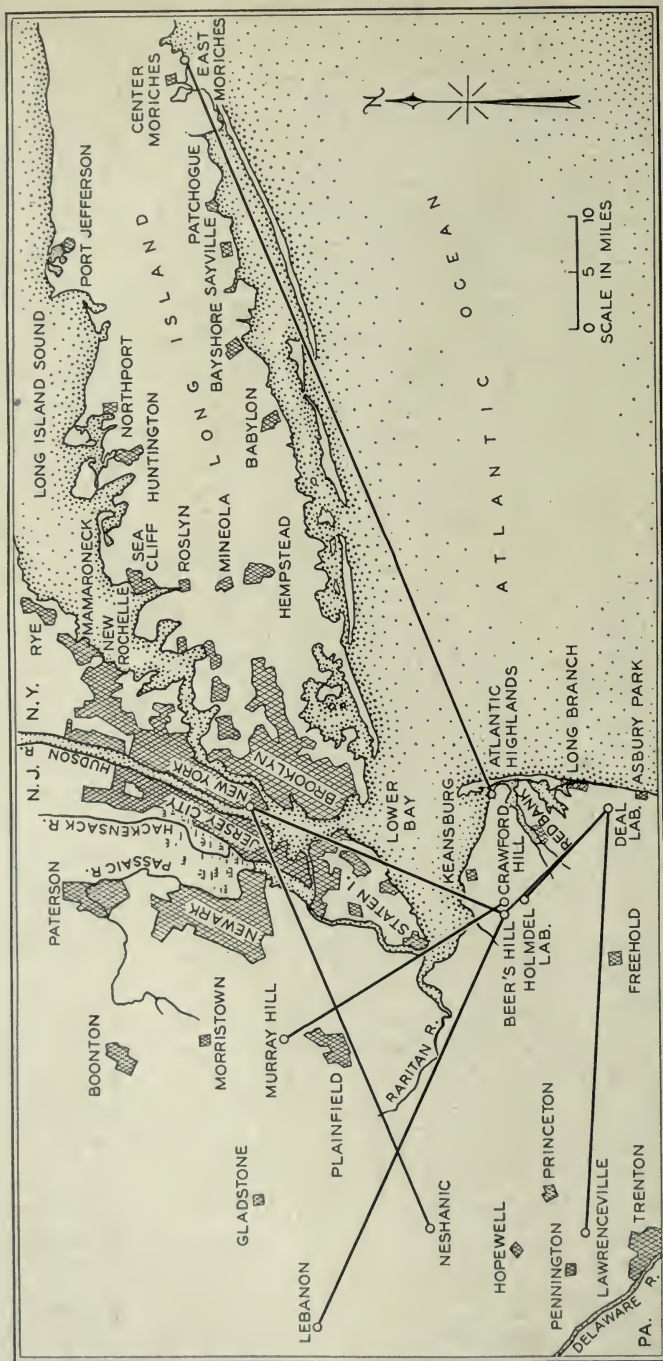


Fig. I-1.—Map showing principal Propagation Paths.

sight is well clear of intervening obstructions. In order to have a rule-of-thumb as to the amount of path clearance desirable, we have suggested that the first Fresnel region should be clear of all obstacles. The first Fresnel region for a given transmitter and receiver is bounded by points for which the length of the path, transmitter to point to receiver, is greater by one-half

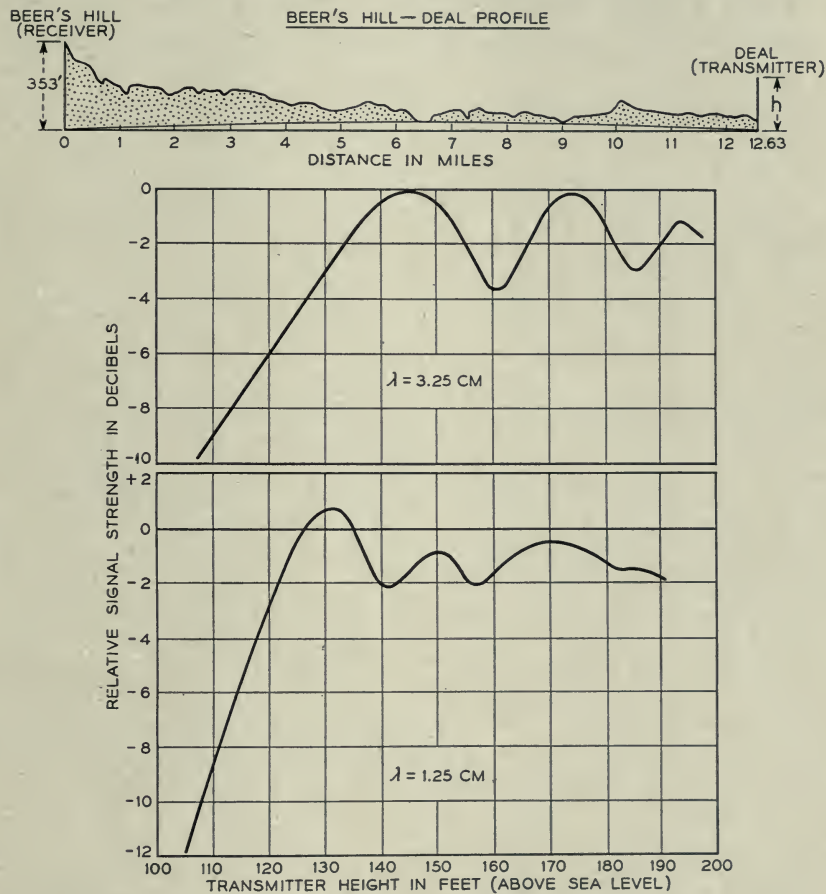


Fig. I-2.—Variable antenna-height tests on Beer's Hill-Deal Path.

wavelength than the direct path from transmitter to receiver; its cross-section by any plane perpendicular to the direct path is the first Fresnel zone in the sense used in optics. A wave can be transmitted with practically no loss through an opening whose area is of the order of the first Fresnel zone. Also, in the case of a smooth reflecting surface between transmitter and receiver, the first Fresnel zone clearance provides a maximum in re-

ceived field strength since the half wavelength path difference plus the 180-degree phase change at reflection causes the direct wave and the reflected wave to arrive in phase at the receiver. In Fig. I-4, the first Fresnel region is sketched on the profile map of a typical microwave link for wavelengths of 3 meters and 3 centimeters.

It should be emphasized that the above remarks on ground reflections apply only for rough terrain and for the case of reflection at a distance from

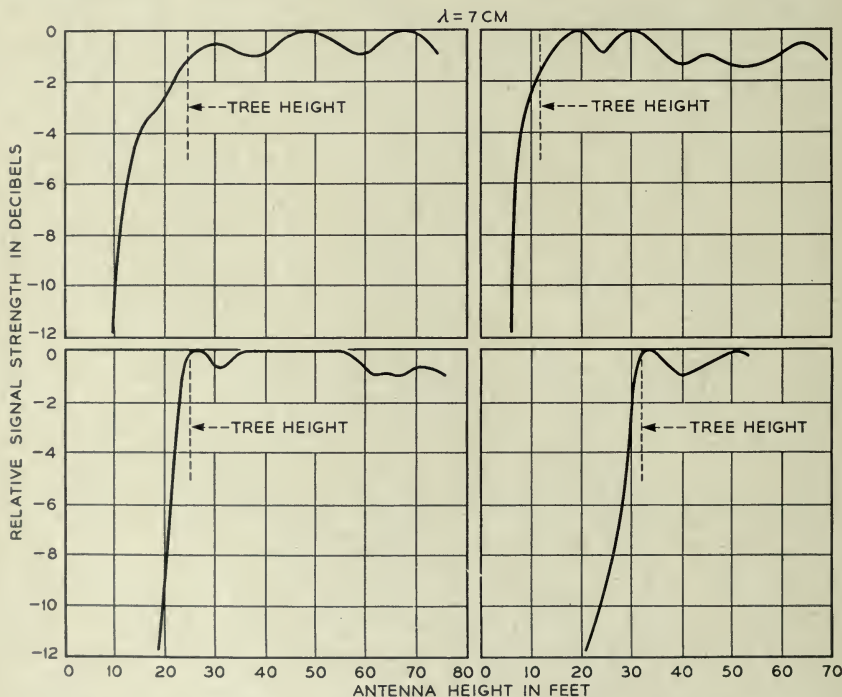


Fig. I-3.—Variable antenna-height tests on several of the New York-Boston repeater circuits. $\lambda = 7 \text{ cm}$.

the terminals. Variable height experiments involving short distances over open flat fields reveal the presence of almost perfect ground reflections at wavelengths as short as 1.25 centimeters. For transmission paths over water, strongly reflected components are often observed. Reports³ of experiments in the Arizona desert indicate a strong ground reflection at wavelength of 3 centimeters. In such locations, and most likely in the plains regions, the presence of substantial ground reflected components may prove to be troublesome.

³ Report No. 6. Electrical Engineering Research Laboratory, The University of Texas, February 1, 1947.

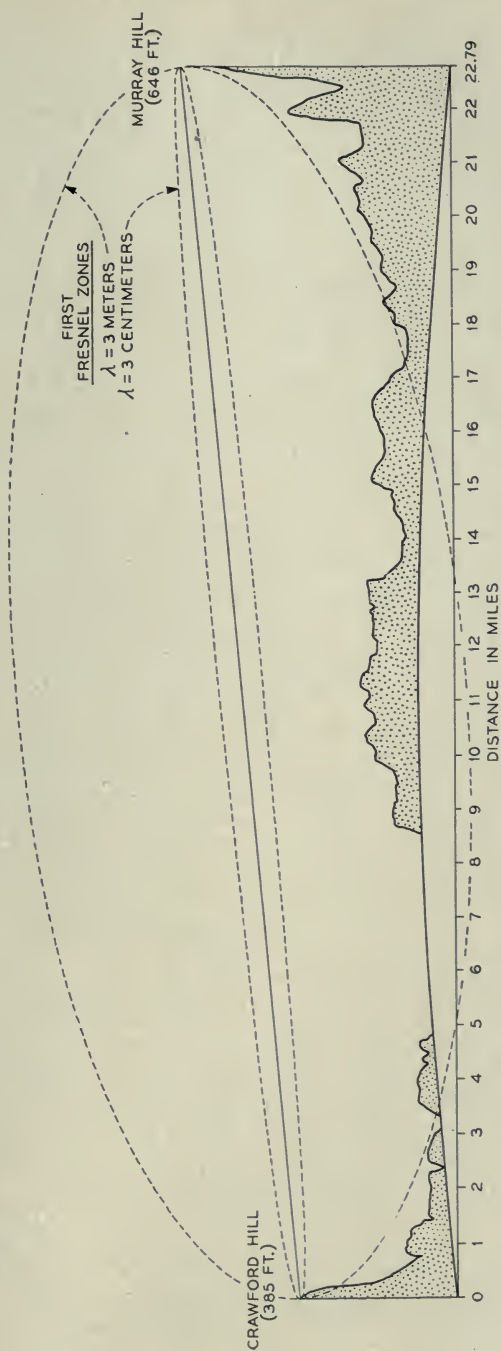


Fig. I-4.—Profile map of Murray Hill-Crawford Hill path showing first Fresnel regions for wavelengths of 3 meters and 3 cm.

ATMOSPHERIC REFRACTION

As in the case of ground reflection, the refractive effect of the atmosphere has been found to play a somewhat varying role, depending upon the wavelength employed. In some of the early work on ultra-short-wave propagation,^{1, 2} the concept of average atmospheric refraction was found to bring about better agreement between observed and calculated results. Due to the variation of temperature and water vapor content of the atmosphere with height above ground, the dielectric constant of the atmosphere normally decreases with height. The effect of this negative dielectric constant gradient is to cause the path of a radio wave to be bent slightly downward toward the earth, thus effectively increasing the horizon distance. It has been suggested that a good approximation for average refraction was to assume the radius of curvature of the ray to be four times that of the earth.¹ This condition is used at the present time to describe a "standard atmosphere."

It was soon found, however, that atmospheric refraction could vary between rather wide limits depending chiefly on the gradient of water vapor with height.⁴ Refraction effects were found to be greater in summer than in winter since the air contains a higher percentage of water vapor in the summertime. A diurnal variation in refraction was also observed on overland transmission paths. During the day, rising convection currents and surface winds, caused by surface heating of the earth, usually produce a well mixed atmosphere near the earth so that "standard" atmospheric conditions prevail. On clear nights, however, particularly if the wind velocity is low, radiation cooling of the earth may cause a temperature inversion in the lower atmosphere; if, also, the water vapor decreases with height, the combined temperature and water vapor effects may add to produce a steep negative gradient in the dielectric constant. Stormy weather and overcast skies usually result in standard atmospheric conditions.

Most of the signal variations observed during a two-year study of propagation of two and four-meter waves over the 39-mile over-land optical path between Beer's Hill, N. J. and Lebanon, N. J.⁵ could be explained satisfactorily on the basis of wave interference between direct and ground-reflected radiations; the relative path lengths, and hence the phases; of these two components of the received field varied with the refractivity of the atmosphere. The fading on the two wavelengths was usually similar in major detail as might be expected from the geometry of the path. On the

^{1, 2} Loc. cit.

⁴ Englund, Crawford and Mumford, "Further Studies of Ultra-Short-Wave Transmission Phenomena", *B. S. T. J.*, vol. 14, pp 369-387; July 1935.

⁵ Englund, Crawford and Mumford, "Ultra-Short-Wave Transmission over a 39 mile 'Optical' Path", *Proc. I. R. E.*, vol. 28, pp. 360-369; August 1940.

few occasions when the fading could not be accounted for in this simple fashion, it was assumed that signal components were arriving from above by virtue of reflections from small, relatively abrupt changes in the dielectric constant of the atmosphere. The existence of such reflections was demonstrated by a frequency-sweep method during propagation studies on the 70 mile over-water path between Highlands, N. J. and East Moriches, Long Island.⁶

With microwaves, where, as stated previously, ground reflections are usually small or absent, one might surmise that changes in atmospheric refraction would have a smaller effect on transmission than at ultra-short-wavelengths where strong ground reflections are present, and that fading should, therefore, be less. Actually the opposite is observed. Fading is found to be more frequent, faster, and deeper as the wavelength is decreased. This frequency effect may be explained in a qualitative fashion by a consideration of the relative sizes of Fresnel zones at ultra-short waves and at microwaves. It is known that the dielectric constant of the atmosphere usually does not vary with height in a smooth linear manner; on calm nights, particularly, very steep gradients in the dielectric constant may exist over small vertical ranges measuring only tens of feet. The effectiveness of these steep gradients would be expected to depend on their extent relative to the size of a Fresnel zone. Thus on a path such as that in Fig. I-4, a steep gradient extending over only a hundred feet would include practically the whole first Fresnel zone at 3 centimeters while it would cover only a small part of a zone at 3 meters wavelength; the effective gradient, therefore, would be considerably less at 3 meters than at 3 centimeters. Analyses based on wave theory show that atmospheric layers, in which the dielectric constant has a steep negative gradient, tend to confine or guide the radiation in much the same way as a waveguide, and that this "trapping" phenomenon, for a given layer thickness, becomes more pronounced as the wavelength is decreased.⁷

The mechanism of microwave propagation is certainly a complicated one, and a considerable amount of experimental work in the fields of radio and meteorology will be required to unravel it. However, it is very difficult to interpret the radio measurements in terms of meteorological data. The chief difficulty is that meteorological measurements often do not give an accurate picture of the atmosphere, particularly at those times when microwave fading indicates that rapid changes of some sort are occurring in the

⁶ Englund, Crawford and Mumford, "Ultra-Short-Wave Transmission and Atmospheric Irregularities", *B. S. T. J.*, vol. 17, pp. 489-519; October 1938.

⁷ H. G. Booker, in England, was the first to call attention to this phenomenon. For more recent work see: C. L. Pekeris, "Wave Theoretical Interpretation of Propagation in Low-Level Ocean Ducts, *Proc. I. R. E.*, vol. 35, pp. 453-462; May, 1947. This paper gives references to other work in this field.

atmosphere. The instruments used to measure temperature and humidity require a few seconds to reach equilibrium—a length of time comparable at times with the period of fading. To measure the variation of dielectric constant with height, the measuring instruments are usually carried aloft by means of captive balloons. A half hour may be required to measure to heights of six or seven hundred feet with the result that the final curve represents an unknown combination of variations of dielectric constant with height and with time. It is also extremely doubtful that the atmosphere is uniform in the horizontal plane—an assumption which is usually made in the theoretical treatment of microwave propagation. It seems likely that the lower atmosphere is far from being a homogeneous fluid but rather may contain small air masses or “boulders” with properties which differ considerably from those of the surrounding air. Reflections from these boulders may be the cause of radar echoes received from the lower atmosphere.⁸ Scintillation fading of microwaves is another evidence of these inhomogeneities in the atmosphere. Scintillation fading, a rapid fluctuation in signal level about a more or less steady average value, increases as the wavelength becomes less and as the path length is increased.

In order to evaluate, on a statistical basis, the effect of atmospheric changes on a typical microwave circuit, extensive measurements of transmission were made over a 40-mile overland path between New York City and Neshanic, New Jersey. The tests covered a period of about two years. Most of the data were obtained at wavelengths of 10, 6.5, and 3.2 centimeters although some data were taken at wavelengths of 42 centimeters and 1.25 centimeters. The results are described in a recent paper.⁹ In many respects, observations were in agreement with those made earlier on the 39-mile Beer's Hill-Lebanon path at wavelengths of 4 and 2 meters and on the 38-mile non-optical path between Deal, N. J. and Lawrenceville, N. J. at a wavelength of 2 meters.¹⁰ The same seasonal and diurnal trends in fading were found; transmission was generally steady during the midday hours and during periods of windy or rainy weather; fading was the same on vertical and horizontal polarizations. However, the character of the fading was different; the fading at microwaves was much faster and deeper than that observed on the ultra-short-wave path. The average daily fading range for July on the New York-Neshanic path was 20 db at 6.5 centimeters compared with a median daily fading range of 8.5 db for 2.0 meters observed in July on the Lebanon-Beer's Hill path.

⁸ H. T. Friis, “Radar Reflections from the Lower Atmosphere”, *Proc. I. R. E.*, vol. 35, pp. 494-495; May 1947 (Correspondence Section).

⁹ A. L. Durkee, “Results of Microwave Propagation Tests on a 40-mile Overland Path”, *Proc. I. R. E.*, vol. 36, No. 2, pp. 197-205, Feb. 1948.

¹⁰ C. R. Burrows, A. Decino and L. E. Hunt, “Stability of Two-Meter Waves”, *Proc. I. R. E.*, vol. 26, pp. 516-528; May 1938.

Other observations on the New York-Neshanic microwave path may be summarized as follows: While all the wavelengths were affected at times of anomalous propagation, the shorter wavelengths faded more severely and the character of the fading was different from that observed at the 42 centimeters wavelength; apparently the 3-10 centimeter range was more sensitive to the fine structure of the atmosphere, as pointed out previously. During non-fading periods, signal levels were very close to the free-space values with the exception of the 1.25 centimeter signal which was usually 15 db or more below the free space value because of atmospheric absorption effects. Some special tests showed that fading was considerably more severe when one of the terminals was lowered so that the transmission path was grazing slightly below line-of-sight. It was also found that fading was about twice as great, in decibels, on the whole path as on either half-section. A statistical analysis, on an hourly basis, of all the data on 6.5 centimeters showed that only one-half of one percent of the total hours had signal minima deeper than 20 db below the free space field. Also during August 1, the day of the most severe fading, the signal was more than 20 db below free space for about one-half of one percent of the time. It was also found that signals of the order of 10 db above free space were equally probable. From a consideration of these statistics, it was decided to engineer the New York-Boston repeater circuit with -20 to $+10$ db allowance for fading on each link.

SPECIALIZED EXPERIMENTS

Much of our more recent work on microwave propagation has been of a specialized nature in which apparatus and experiments have been designed more for the purpose of studying the mechanism of anomalous propagation than for making a statistical analysis of the transmission. Perhaps the most informative experiments have been those in which narrow beam scanning antennas were used to explore the incident wave fronts.

The first of these antennas had an aperture of 20 feet and a beam width between half-power points of $\frac{1}{2}$ degree at the design wavelength of 3.25 centimeters. It was built for the purpose of establishing a practical limit to the size, and hence the directivity, of microwave repeater antennas from the standpoint of variations in the angle of arrival of the received wave. It had been realized, of course, that variations in the refractivity of the atmosphere would cause some deviations in the path of the wave. While these deviations should be negligible in comparison with the beam width of antennas normally used in the ultra-short-wave region, they might be comparable with the beam widths readily obtainable in the microwave region.

Using this antenna for measurements in the vertical plane and another identical antenna for measurements in the horizontal plane, angle-of-arrival data were obtained during the summer of 1944 over a twenty-four mile, partly over-water, path between Beer's Hill, N. J. and New York and over a thirteen mile over-land path between Beer's Hill, N. J. and Deal, N. J.¹¹ In the horizontal plane, deviations in the angle of arrival were rather uncommon and were not greater than ± 0.1 degree from the true bearing of the transmitter. In the vertical plane, angles of arrival above the true elevation of the transmitter were observed to be as much as 0.5 degree on the New York path and 0.3 degree on the Deal path during times of anomalous propagation. From these measurements it was concluded that microwave repeater antennas could be made highly directive in the horizontal plane but should have beam widths somewhat greater than $\frac{1}{2}$ degree in the vertical plane unless means for steering the beams are provided.

Although the $\frac{1}{3}$ degree beam width of these scanning antennas was sharp enough to permit the separation of the direct and the water-reflected components on the New York path, and to demonstrate the anomalous behavior of each, there was evidence that occasionally there were signal components so close together in angle that a sharper antenna would be required to resolve them. Consequently a scanning antenna of the metal-lens type was constructed for operation at 1.25 centimeters. The aperture of this antenna was 20 feet in the long dimension; the beam width was 0.12 degrees. Using this antenna and also the 3.25 centimeter scanning antennas, angle-of-arrival measurements were made in the summer of 1945 on the Deal-Beer's Hill path.¹² The most noteworthy result of these observations was the demonstration of multiple-path transmission. Two, three and, at times, four distinct signal components were observed simultaneously during one night when the transmission was extremely disturbed. These transmission paths generally were above the true direction of the transmitter; at one time, a weak signal was arriving at an angle of 0.75 degree relative to the line of sight. These components varied in angle of arrival and in signal amplitude. Wave interference among them caused severe fading on broad beam antennas that would accept all the wave paths.

Another significant result of these angle-of-arrival measurements was evidence that the transmission mechanism was very similar for wavelengths of 3.25 and 1.25 centimeters. Angles of arrival, measured simultaneously at the two wavelengths, agreed very well for times of single-path transmission; multiple-path transmission was observed on both wavelengths although the 3.25 centimeter antenna was too broad to resolve the com-

¹¹ W. M. Sharpless, "Measurement of the Angle of Arrival of Microwaves", *Proc. I. R. E.*, vol. 34, pp. 837-845; November 1946.

¹² A. B. Crawford and W. M. Sharpless, "Further Observations of the Angle of Arrival of Microwaves", *Proc. I. R. E.*, vol. 34, pp. 845-848; November, 1946.

ponents completely. This result suggested that the 1.25-centimeter scanning antenna might be a very useful tool for investigating the fading mechanism at 7 centimeters wavelength.

The 22.8-mile path between Crawford Hill and a hill on the Murray Hill Laboratory property was chosen for study as a representative link in a repeater circuit. (See Profile Map, Fig. I-4.) Transmitters for the 1.25-centimeter and 7 centimeter wavelengths were installed in the 100-foot tower at Murray Hill. At the Crawford Hill receiving site were the narrow-beam scanning antenna and a broad beam antenna for 1.25-centimeter operation; also two broad beam antennas, spaced vertically 15 feet, for 7-centimeter operation. In addition, a 1.25-centimeter radar could be operated with the scanning antenna. A corner reflector target, $5\frac{1}{2}$ feet on a side, was located at the Murray Hill tower. The signal reflected by this target was about 10 db stronger than the spurious reflections from other objects at the same range. By making use of this target and ground echoes at intermediate distances, the radar technique provided a considerable amount of useful information concerning the transmission phenomena.

Measurements were made on this path during the summer of 1946. As had been hoped, the observations showed that transmission on 1.25 centimeters and 7 centimeters was often affected by the same conditions except, of course, for atmospheric absorption effects at the 1.25-centimeter wavelength. While it was not possible to arrive at explanations for all the fading observed, the deep minima in the 7-centimeter signal, i.e., fades to levels of 15 to 20 db or more below the free space field, usually were the result of one of three types of propagation*:

Type 1. The 7-centimeter fading was of the rapid, large amplitude type characteristic of wave interference. The 1.25-centimeter scanning records showed the presence of multiple-path transmission in which two or more readily separable wave paths were observed. While the signals on both of the vertically-spaced 7-centimeter antennas faded about the same in amplitude, their signal minima did not occur simultaneously. A space diversity system would be successful in reducing the effects of this type of fading.

Type 2. The 7-centimeter fading was somewhat slower than in Type 1, but still had the appearance of wave interference. The 1.25-centimeter scanning records appeared to be of the single path variety. However, close inspection showed that, in all probability, more than one transmission path was involved but the 0.12 degree beam of the antenna was not sharp enough to resolve them. The signals received on the vertically spaced 7-centimeter antennas faded together so that space diversity would not be expected to be successful unless an extremely large spacing of antennas were used.

* Recently, on a different overland path having barely one Fresnel zone clearance, an important fourth type has been observed when atmospheric refraction gives the ray path a curvature opposite to that of the earth, thus effectively reducing the path clearance.

Type 3. The 7-centimeter signal would fade to a low level and remain there for a considerable period of time; sometimes for an hour or so. The character of the fading was unlike that caused by wave interference. The 1.25-centimeter signal was simultaneously at a low level and the scanning records showed that only one path was involved. Reception was almost

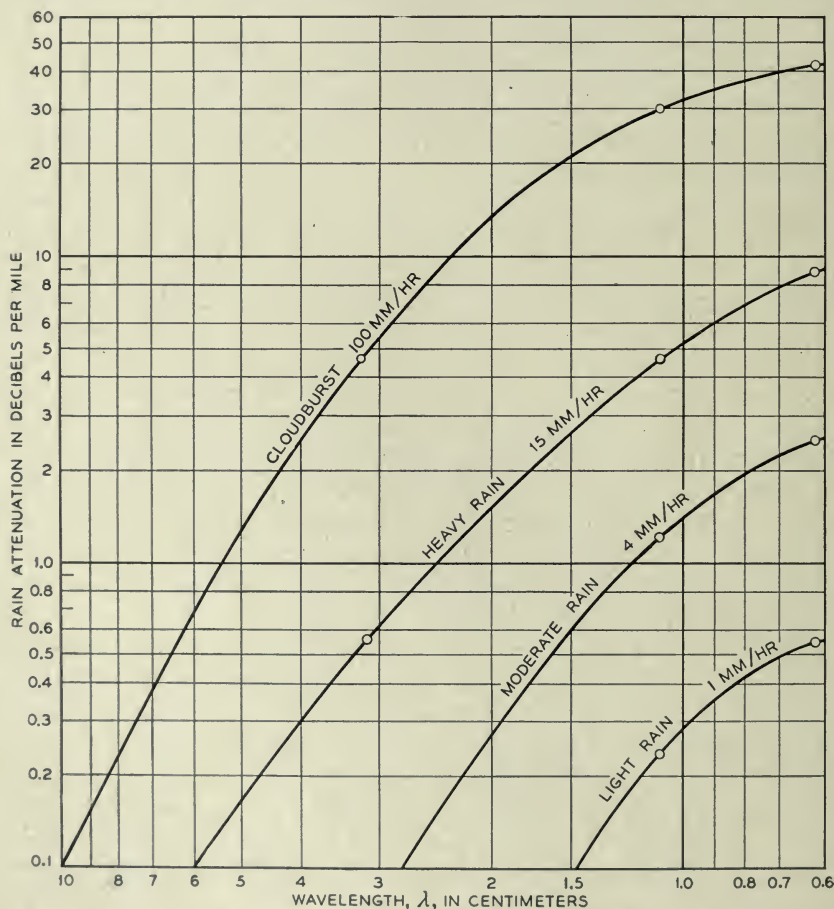


Fig. I-5.—Rain attenuation vs. wavelength.

identical on the two vertically-spaced 7-centimeter antennas. Radar observations suggested that this type of fading was due to attenuation by a reflecting layer in the atmosphere at a height intermediate to the heights of the transmitters and receivers. It was observed, for example, that while the echo from the Murray Hill corner reflector was absent, strong echoes were received from the hill directly in front of Murray Hill and some 250

feet lower in height; sometimes multiple paths were observed with this echo. Space diversity would fail to improve transmission under these propagation conditions, and no other means of improvement is apparent except, perhaps, an alternate path. Fortunately, this type of fading was the least frequent of the three types which were characterized by low signal levels.

RAIN ATTENUATION AND ATMOSPHERIC ABSORPTION

Attenuation effects due to rainfall and absorption by atmospheric gases become increasingly important at the short-wave end of the microwave region. Measurements of rain attenuation have been made at the Holmdel Laboratory^{13, 14}; the results are summarized in Fig. I-5. These curves show that for wavelengths above about 5 centimeters, rain attenuation is not very serious except for rains of cloudburst proportions. However, at wavelengths of one centimeter and less, even moderate rainfall will cause large attenuations on paths of the order of 10-20 miles in length.

Absorption by atmospheric gases, principally water vapor and oxygen, becomes important at wavelengths below about 1.5 centimeters. According to the theoretical work of Dr. J. H. Van Vleck, Harvard University, water vapor has an absorption band at 1.33 centimeters and oxygen has bands at 0.5 and 0.25 centimeters. Measurements made on the Deal-Holmdel path at 1.25 centimeters were in fair agreement with Van Vleck's results and indicated that a typical value of atmospheric absorption for this locality in summertime is about 0.4 db per mile.¹⁴

SUMMARY

In the ultra-short-wave region, transmission has been found to be affected mainly by ground reflections and variable atmospheric refraction; only occasionally are atmospheric reflecting layers and trapping phenomena involved. These wavelengths ordinarily are not transmitted to great distances along the surface of the earth, but are diffracted around obstacles. They are used for local broadcasting and mobile radio communication.

Microwaves are attractive for radio repeater circuits since they permit the use of wide transmission bands. Ground reflections are apparently of small importance with terrain such as that of the Eastern seaboard and substantially free-space propagation is obtained during non-fading periods over optical paths which have approximately "first Fresnel region" clearance. Atmospheric reflecting layers and trapping phenomena are frequently observed and signal variations are considerably greater than in the ultra-

¹³ Sloan D. Robertson and Archie P. King, "The Effect of Rain upon the Propagation of Waves in the 1- and 3-Centimeter Regions", *Proc. I. R. E.*, vol. 34, pp. 178P-180P; April 1946.

¹⁴ G. E. Mueller, "Propagation of 6-Millimeter Waves", *Proc. I. R. E.*, vol. 34, pp. 181P-183P; April 1946.

short-wave region. Although fading becomes worse as the wavelength is decreased, the advantages of increased antenna gain and directivity possible at the shorter wavelengths suggest the use of a wavelength just above the region where rain attenuation becomes objectionable; i.e. above about 5 centimeters.

The use of two antennas, operated in space diversity, should reduce the effects of fading caused by multiple-path transmission. The use of space diversity may be essential in those localities where strong ground reflections are present. On the basis of the comparatively weak ground reflections measured on the New York-Boston path it was decided to avoid the complications that would result from the use of space diversity in this experimental system.

II. REPEATER CIRCUIT PLANNING

The diagram in Fig. II-1 shows schematically a repeater circuit. At the input terminal toward the left the signal, S , is fed to the terminal's transmitting antenna. The radiated signal is propagated as discussed in Section I and produces the signal power s_1 in the output of the receiving antenna of repeater 1. The signal is then amplified G_1 times and radiated toward repeater 2 and this process is repeated until the signal finally appears in the output terminal towards the right. In each repeater the signal is gain-controlled automatically for the same level of output powers, i.e., $S = S_1 = S_2 = \dots = S_n$. It is assumed that the signals are amplitude or frequency-modulated C.W. carriers of substantially the same frequency in each link and that the repeaters have linear amplifiers. The diagram shows a West-to-East circuit only. A circuit for the opposite direction requires duplication of all the equipment with the exception of the antenna supporting towers.

Some simple formulas for the repeater gain and the signal-to-noise ratio at the terminal will be given in this section, without going into any details on propagation phenomena, antennas, amplifiers, etc. The formulas will orient the reader in regard to the importance of quantities such as:

$$\left. \begin{array}{l} d = \text{Repeater separation} \\ \lambda = \text{Wavelength of signal} \\ A = \text{Effective area of each antenna} \\ F = \text{Noise figure of each repeater amplifier} \\ B = \text{Bandwidth of circuit in cycles/sec.} \end{array} \right\} \text{same units of length}$$

The free space attenuation (S_{x-1}/s_x) of link "x" is¹⁵

$$L_x = \frac{d_x^2 \Lambda^2}{A^2}$$

¹⁵ H. T. Friis, "A Note on a Simple Transmission Formula", *Proc. I. R. E.*, vol. 34, No. 5, pp. 254-256; May 1946.

Allowing the signal power to fade by a factor M below this value, the maximum gain of repeater "x" must be

$$G_x = \frac{d_x^2 \lambda^2}{A^2} M_x \quad (\text{II-1})$$

The total maximum gain in the circuit is $G_T = G_1 \times G_2 \times \cdots G_n$ which for the same repeater spacings and fading allowances becomes

$$G_T = G^n = \left(\frac{d^2 \lambda^2 M}{A^2} \right)^n = \left(\frac{d^2 \lambda^2 M}{A^2} \right)^{D/d} \quad (\text{II-2})$$

where D is the total length of the circuit. Because of distortion, original costs, and maintenance costs, this total gain should be made as small as possible.

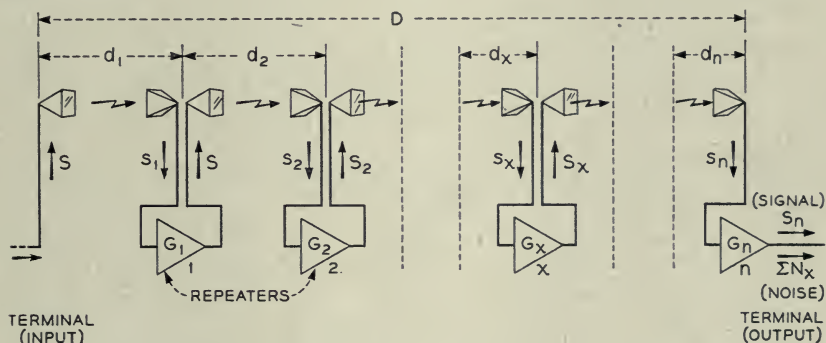


Fig. II-1.—Repeater circuit with n links.

The following example illustrates the maximum gain required of the amplifier in a repeater:

For $d = 4 \times 10^4$ meters (25 miles), $\lambda = 0.075$ meters ($f = 4000$ megacycles), $A = 4.6$ meter² (50 sq. ft.) and $M = 100$ (20 db), we have

$$G = 4.3 \times 10^7 \text{ (76 db)}$$

The noise output of a repeater due to noise sources in the repeater itself is approximately¹⁶

$$N_x = 4 \times 10^{-21} F B G_x \text{ Watts}$$

or from (II-1)

$$N_x = 4 \times 10^{-21} F B \frac{d_x^2 \lambda^2}{A^2} M_x$$

¹⁶ H. T. Friis, "Noise Figures of Radio Receivers", *Proc. I. R. E.*, vol. 32, pp. 419-22; July 1944.

This noise power is transmitted without gain or loss to the output terminals of the repeater circuit. The total noise power at the output terminals is therefore

$$\Sigma N_x = 4 \times 10^{-21} FB \frac{\lambda^2}{A^2} \Sigma (d_x^2 M_x)$$

Assuming the same repeater spacings and fading factors M in each link and substituting D for nd ,

$$\Sigma N = 4 \times 10^{-21} FBD \frac{d\lambda^2}{A^2} M \text{ Watts} \quad (\text{II-3})$$

The signal-to-noise ratio at the output terminals is $S/\Sigma N$. The circuit should be designed for a signal power, S , as low as possible. Therefore, it is very important to choose values of the several factors in (II-3) which give a low level of output noise.

Assuming a noise figure $F = 20$ (13 db) and bandwidth $B = 10$ megacycles, eight repeaters of the type described in the above example will have,

$$\Sigma N = 2.8 \times 10^{-4} \text{ Watts}$$

or, assuming a required minimum output signal to noise ratio of 30 db, the output power must be $S \geq 0.28$ Watts.

In this example it has been assumed that the signals in all links have faded simultaneously 20 db below the free space value, which may only happen a fraction of a percent of the time. Most of the time the signal-to-noise ratio will be higher than the assumed 30 db and under normal transmission conditions it will be 50 db (fading allowance factor $M = 1$).

Assuming the same repeater spacings and fading allowances in each link, equation (II-1) and (II-3) give the following formula for the ratio of the output power to noise figure of the repeater amplifier

$$S/F = 4 \times 10^{-21} G B (S/\Sigma N)n \quad (\text{II-4})$$

Equations (II-1) and (II-4) are the important design equations for the repeater amplifier.

The factors in (II-2) and (II-3) will now be discussed briefly. (II-3) shows that the noise figure F should be as small as possible. If by improving the equipment the noise figure could be halved, then the signal power S could also be halved (unless interference from other microwave circuits predominate). Later on the noise figure will be discussed further.

The bandwidth B is determined by the characteristics of the signal it is desired to transmit and by the method of transmission. Our aim has been to provide 10-megacycle bands which are sufficient for transmission of standard television signals by AM or low index FM.

An increase in the effective area, A , of the antennas reduces both the total gain required of the amplifier and the output noise power. Crosstalk between the several antennas in a repeater station and interference from outside sources also decrease as the antennas are increased in size because of the increased directivity. Therefore, the antennas should be as large as maintenance and initial costs will permit. Antennas will be discussed in detail in Section III.

Equations (II-2) and (II-3) show that the wavelength λ should be small. Also more frequency space or signal channels may be had at shorter wavelengths. On the other hand, the fading factor M increases somewhat as the wavelength is decreased and, besides, attenuation due to rain sets a lower limit for λ in the region of 5 centimeters. The status of the apparatus art has also been an important factor, but it now permits utilization of the wavelength range extending upward from 3 centimeters. Since the war, our work has been concentrated on a 10% band around 4000 megacycles or $7\frac{1}{2}$ centimeter wavelength. The manner in which this 4000-megacycle band may be divided up into separate channels is explained in Section IV.

The effects of varying the repeater separation d will now be discussed. d appears in the denominator of the exponent of (II-2) which indicates that large separations are favorable, while (II-3) shows that a decrease in separation cuts down the noise. Small separations are very costly, the cost being almost inversely proportional to the separation. We have concluded from propagation studies and site surveys that in the eastern part of the United States it is desirable to use separations of about 30 miles, which generally provide line-of-sight paths with reasonable tower heights. It should be mentioned that the fading allowance factor M is not independent of d ; an increased d requires a larger fading factor.

III. ANTENNA RESEARCH*

There are three electrical characteristics which repeater antennas should possess. The first is high gain (large effective area), as this will reduce the path loss and accordingly the requirements on transmitter power. The second is good directional qualities so as to minimize interference from outside sources and also interference between adjacent antennas. The third is a good impedance match so that reflections between the antenna and the repeater equipment will not distort the transmitted signals. These characteristics should preferably be attainable without the imposition of severe mechanical or constructional requirements.

It was felt that a 10-foot round or square antenna would be the largest that maintenance and initial cost would permit. Propagation studies also

* This section prepared by W. E. Kock who performed the major part of the work on antennas.

showed that the variations in angle of arrival of the distant signals would be small compared to the beam width of 10-foot antennas. First experiments were therefore made with 10-foot diameter parabolic "dish" type antennas. The experimental models were made of wood with a metallized reflecting surface consisting of silver conducting paint. Fairly satisfactory tolerances were met in these first models, but it was anticipated that trouble would be experienced in constructing a permanent metal paraboloid of that size to the required tolerances without the use of a heavy and costly supporting means for the parabolic sheet. It was also found that an ice coating a quarter wavelength thick on the reflecting surface, when wet, acted as an effective absorber of power,[†] since the sheet of water is resistive and is backed up by the reflector. Such a condition could produce an intolerable drop in received signal and would have to be prevented by providing the dish with a plastic cover. As this cover should preferably house the feed also, it would have presented a difficult supporting problem.

Two electrical shortcomings of the paraboloid antenna also presented themselves. First, it was found extremely difficult to obtain a satisfactory impedance match between the antenna and the feed line. This was true partly because of energy reflected from the dish re-entering the feed horn, (this produced a constant 0.6 db standing wave ratio in the feed line), and partly because of the problem of matching the feed horn itself over the desired 400 megacycle band. Secondly, the mutual interference or "crosstalk" between two paraboloids was found to be only 50 to 60 db down when placed back to back** (Fig. III-1).

A type of reflector antenna was later investigated, which, although larger physically than a dish having the same aperture area, overcomes the above two objections.¹⁷ It is shown sketched in Fig. III-2. The photograph of Fig. III-3 shows the antenna lying on its side. It can be seen that the feed is effectively "offset" and reflection back toward the feed is eliminated; the experimental model of Fig. III-3 showed only 0.1 db standing wave ratio in the feed line over a 10% band of frequencies. Furthermore, a horn or "shielded" type feed is used which confines the energy and minimizes stray radiation, and measurements indicate that the back-to-back crosstalk suppression of two such antennas will be high. This long horn is also partly responsible for the excellent impedance match. A horn having a large aperture "matches" free space quite well and the slight mismatch at the throat can be tuned out over a wide band of frequencies. This is not true

[†] A waveguide termination in common use today employs a resistive sheet placed one-quarter wavelength in front of a conducting plate; this device absorbs practically all the power falling on it.

** Back to back crosstalk suppression in the order of 125 db would be desirable for repeaters receiving and transmitting on the same frequency.

¹⁷ U. S. Patent #2,236,393, H. T. Friis and A. C. Beck.



Fig. III-1.—Measuring the back to back crosstalk of two 10 ft. paraboloid antennas.

of the short, small aperture horn used in feeding the dish antenna. The antenna displayed an effective area which was 66% of its actual area which is only 0.9 db below the theoretical maximum.

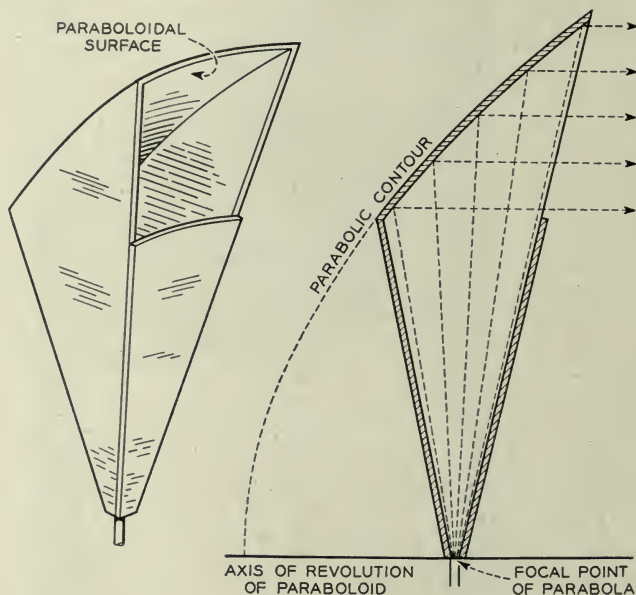


Fig. III-2.—Schematic of horn-reflector antenna.

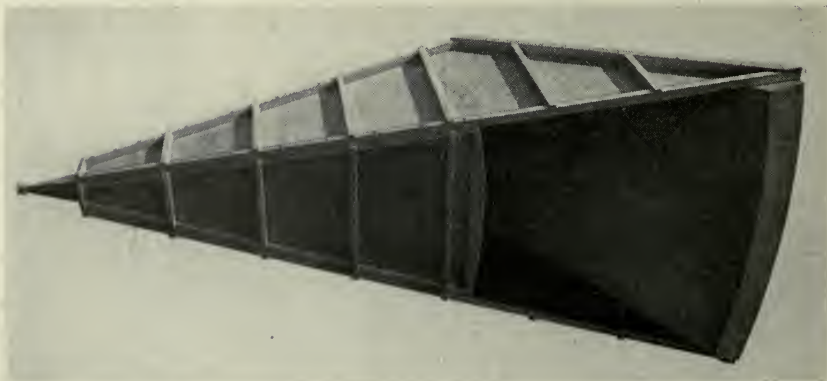


Fig. III-3.—Experimental model of horn reflector antenna.

The expected gain and directional characteristics of an antenna can be realized only if the emerging wave fronts are truly plane. Since deviations greater than $\pm \frac{1}{16}$ wavelength can materially impair the antenna perform-

ance,¹⁸ reflector antennas have difficult tolerance requirements imposed upon them. For example, at 4000 megacycles, the 10-foot reflector must conform to parabolic shape to within $\pm \frac{1}{8}$ inches, and any twist or warp

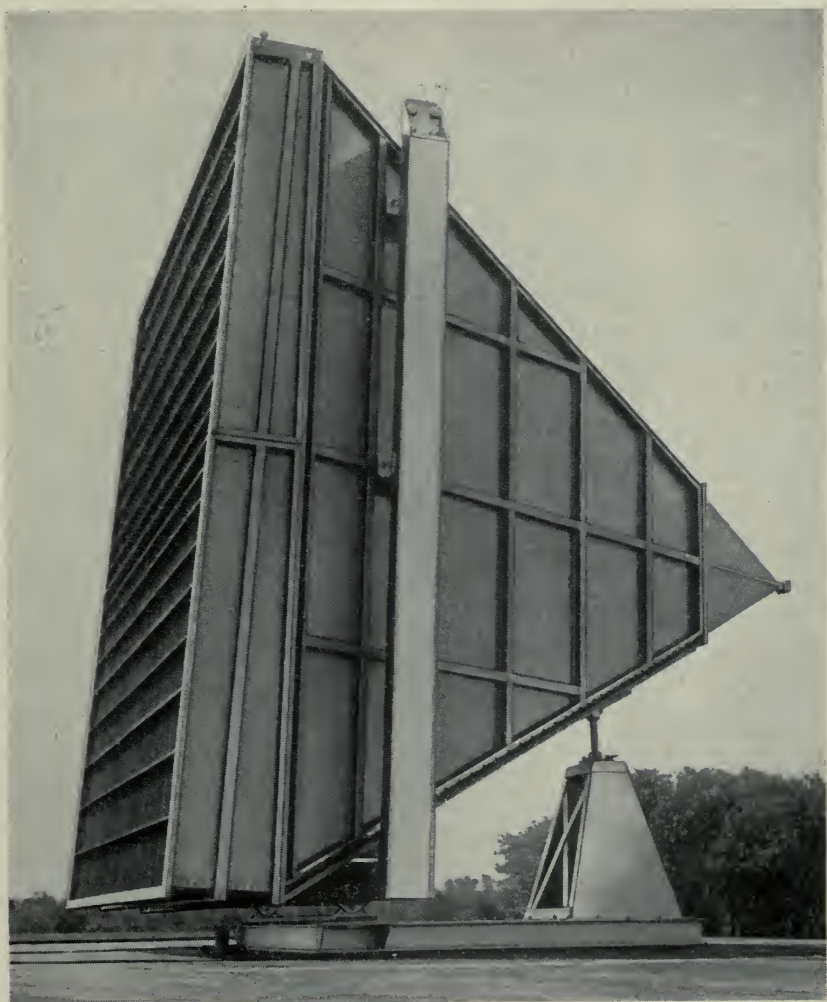


Fig. III-4.—Shielded metallic lens antenna.

of the reflector greater than this would be objectionable. Lenses, however, possess the property that a twist or warp in them does not impair their

¹⁸ See, for example, "Radar Antennas", H. T. Friis and W. D. Lewis, *B. S. T. J.*, vol. 26, p. 219, April 1947, Figs. 17 and 28.

beam-forming properties, and, with the development of metal lenses for microwaves,¹⁹ this type of antenna appeared to lend itself very well to repeater applications. The "shielded" type lens, which is a lens in the mouth of a short horn, is shown in Fig. III-4. This antenna, which was developed for the New York-Boston circuit,* possesses the property of excellent

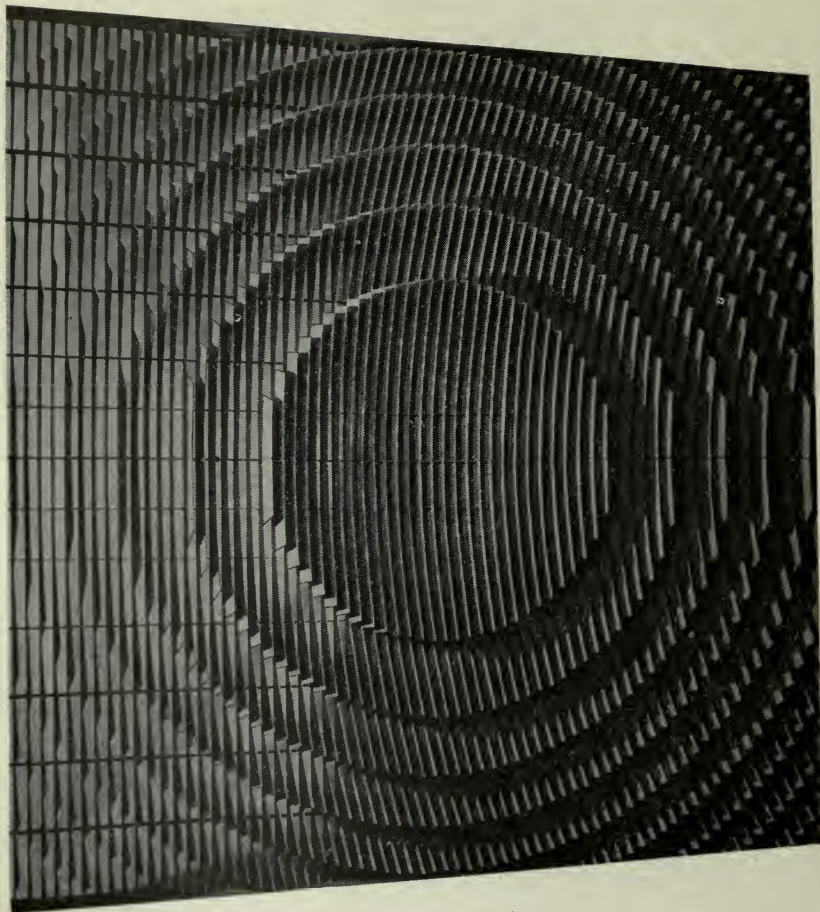


Fig. III-5.—Internal view of lens for the shielded metallic lens antenna.

crosstalk suppression both side to side (85 db) and back to back (125 db). Within the horn, the small amount of energy reflected back from the lens is directed away from the feed by tilting the lens, a procedure which does not noticeably affect the radiation characteristics, but which results in a fairly

¹⁹ W. E. Kock, Metal Lens Antenna, *Proc. I. R. E.*, vol. 34, p. 828, November 1946.

* Developed by W. E. Kock and R. W. Friis.

good impedance match (under 0.8 db standing wave ratio) over the desired 400-megacycle band of frequencies. The lens in the mouth of the horn also provides a convenient support for a plastic impregnated Fiberglass sheet which acts as a weatherproof cover and protects the lens against ice forming between the plates.

The lens itself, Fig. III-5, is based on waveguide principles and causes the wave to be refracted by virtue of the fact that waves confined between plates parallel to the electric vector acquire a phase velocity higher than their free space velocity in accordance with the equation:

$$v_{\text{lens}} = v_{\text{free space}} / \sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}, \quad (\text{III-1})$$

where λ is the wavelength and, a , the distance between the plates. The index of refraction is thus less than one, and a converging lens must be made concave.

As seen in Fig. III-5, the lens is stepped to reduce its thickness. As a consequence of this stepping, the efficiency at midband of the antenna (50%), is a good deal less than the theoretical value of 81%. Furthermore, the index of refraction varies with wavelength, as seen from equation III-1, and this results in a defocussing of the lens, with a consequent drop in gain, at wavelengths different from the design wavelength. This amounts to a drop in gain of 1.5 db at the edges of a 400 megacycle band; however, its other characteristics of impedance, side lobe suppression (70 db in the two rear quadrants), crosstalk, and ease of construction, help to make up for the gain deficiency.

Measurements taken on the antenna when a thick coating of ice had formed on the plastic cover indicated that the impedance match was impaired (the maximum standing wave ratio increased from .8 db to 1.6 db), but that the gain was not appreciably affected (less than 1 db). Since propagation experiments indicate that severe atmospheric fades are not likely to occur during the winter months, some of the fading allowance can be applied against ice loss.

There was some doubt that the crosstalk figures quoted above could be relied upon during heavy rainfalls, as there was indication that the signal transmitted from one antenna might be reflected from the rain and thus caused to enter an adjacent side-by-side antenna. Measurements during a moderately heavy rain proved that this effect was small, but large enough so that the 85 db side-to-side figure was approaching a limit for the 4000 megacycle band.

The measurement of antenna characteristics involves microwave techniques whose development is an important part of a research program.

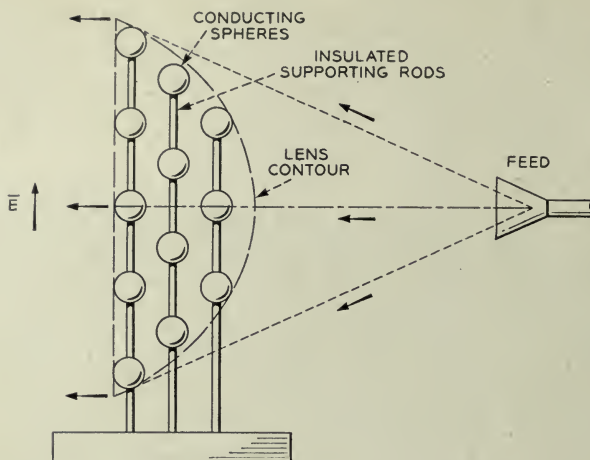


Fig. III-6.—Schematic of metallic delay lens using metal balls as the delay elements.

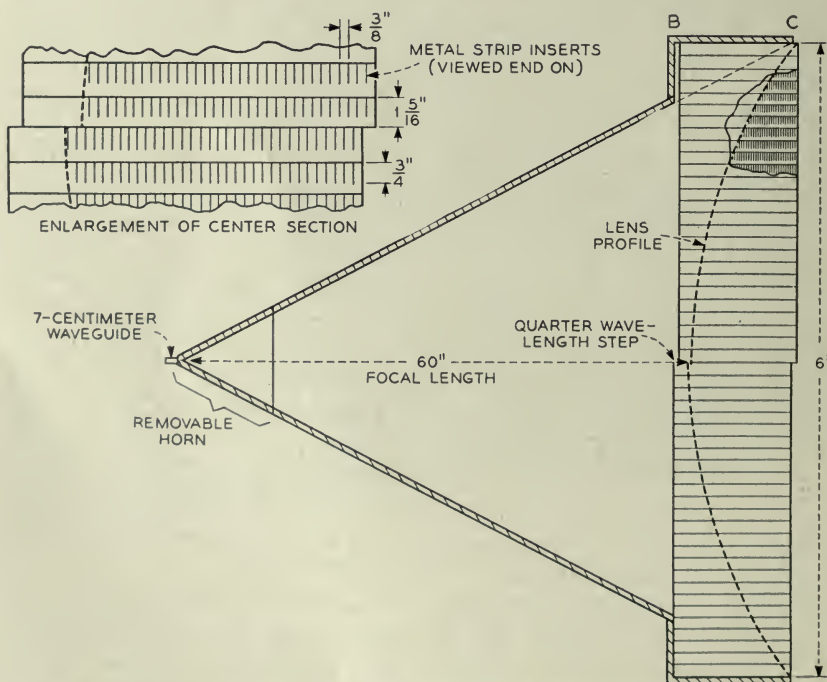


Fig. III-7.—Schematic of metallic delay lens for repeater applications using metal strips as the delay elements.

The antenna measuring methods which were employed in our repeater research follow along the lines of those described in a recent paper.²⁰ The very large signal ratios of 120 db or more necessitated double detection receivers and low noise figure crystal converters. Pattern and gain measurements of the antennas required measuring sites having large unob-



Fig. III-8.—A view of the partly assembled lens of Fig. III-7.

structed areas; these measurements were conveniently taken at the Holmdel Radio Research Laboratory. Impedance measurements involved the usual microwave equipment such as standing wave detectors in waveguide form, signal generators and calibrated receivers.

Research is now underway on an improved metal lens with gain and bandwidth properties which are superior to the lens of Fig. III-5. These lenses²¹

²⁰ C. C. Cutler, A. P. King, W. E. Kock, "Microwave Antenna Measurements", *Proc. I. R. E.*, vol. 35, No. 12, pp. 1462-1471, December 1947.

²¹ Winston E. Kock, "Metallic Delay Lenses", *B. S. T. J.*, vol. 27, pp. 58-82, January 1948.

employ an artificial dielectric material which duplicates, on a much larger scale, processes occurring in a true dielectric. This involves arranging conducting elements in a three dimensional array or lattice structure to simulate the crystalline lattice of the true dielectric. Such an array responds to radio waves just as a molecular lattice responds to light waves, and if the spacing and size of the elements is small compared to the wavelength, the index of refraction is substantially constant, so that lenses made of this material are effective over large wavelength bands.

A lens employing conducting spheres as the lattice elements is sketched in Fig. III-6. A more convenient structure for large lenses is shown in Fig. III-7 and III-8; it uses thin metallic strips, with the width dimension parallel to the electric vector. Slotted polystyrene foam sheets support the strips and they are stacked up to form the lens. A quarter wavelength step in the lens causes the reflections from the lens surfaces to cancel at the feed point, which, in the drawing, is the apex of the horn shield.

Over a 10% wavelength band, a 6 foot square shielded lens antenna of this type exhibited an efficiency of better than 60% and the impedance mismatch due to the lens produces only a 0.2 db standing wave ratio in the feed line. This antenna thus retains the dimensional tolerance, weight, size and crosstalk advantages of the shielded lens over the shielded reflector, and has the advantage of higher gain and broader band performance over the shielded metal plate lens.

IV. FILTER RESEARCH*

Frequency space for common carrier radio relay systems is available in blocks several hundred megacycles wide. Where heavy traffic is to be carried such bands must be efficiently exploited. This may in time be accomplished by using extremely wide band amplifiers, for example traveling wave tubes; however, more immediate success is offered by the possibility of operating a number of narrower band circuits of different frequencies. This could be done by using a separate transmitting and receiving antenna for each circuit. But each antenna, for sound technical reasons, must be large and expensive and in addition requires adequate tower support. Consequently there is a need for filters which can connect a number of individual radio channels to a common antenna.

The design²² of these radio frequency branching filters must be coordinated with the design of the relay system as a whole. At lower frequencies where little or no antenna crosstalk protection can be counted on it is natural

* This section prepared by W. D. Lewis and L. C. Tillotson who were responsible for a large part of the research on filters.

²² For more detailed discussion see, W. D. Lewis and L. C. Tillotson "A Constant Resistance Branching Filter for Microwaves," *B. S. T. J.*, vol. 27, pp. 83-95, Jan. 1948.

to lump transmitting frequencies in one group and receiving frequencies in another. When separate microwave shielded lens antennas are employed for transmitting and receiving in each direction it becomes practical to use a frequency plan in which transmitting and receiving frequencies are interleaved. Such a plan eases filter requirements considerably and has

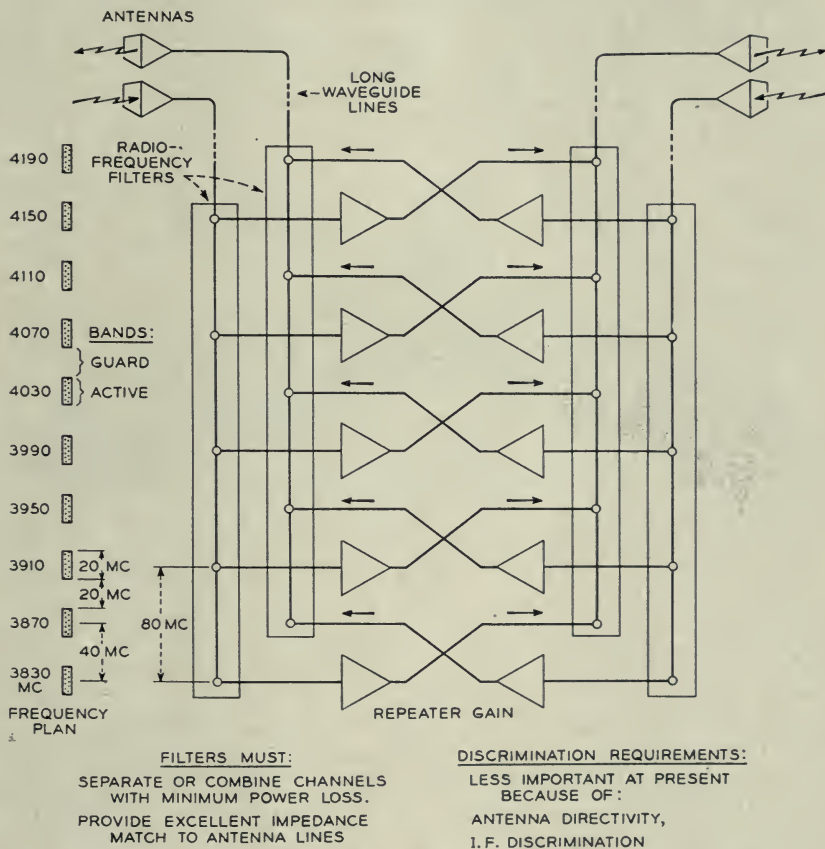


Fig. IV-1.—Schematic diagram of a possible five channel radio repeater station.

certain other advantages to be discussed in Section V. A possible repeater employing such a frequency scheme is illustrated schematically in Fig. IV-1.

If a radio frequency branching filter is to fit properly into a repeater it must separate or combine channels without excessive loss of signal. In addition it must provide an excellent match to the long line which leads to the antenna, otherwise troublesome echoes in this line may be caused.

Because of IF amplifier band-pass characteristics, suppression requirements on the filter, except possibly in the vicinity of receiver image bands, are not large.

Microwave band-pass filters consisting of one or more cavities arranged in sequence along a waveguide have been known for some time. The frequency, bandwidth and discrimination characteristics of such filters can all be chosen within wide limits by appropriate design of the cavities and the means for coupling to them. These filters are analogous to lumped-circuit channel-passing filters and can in principle, like them, be connected in groups to provide a branching network.

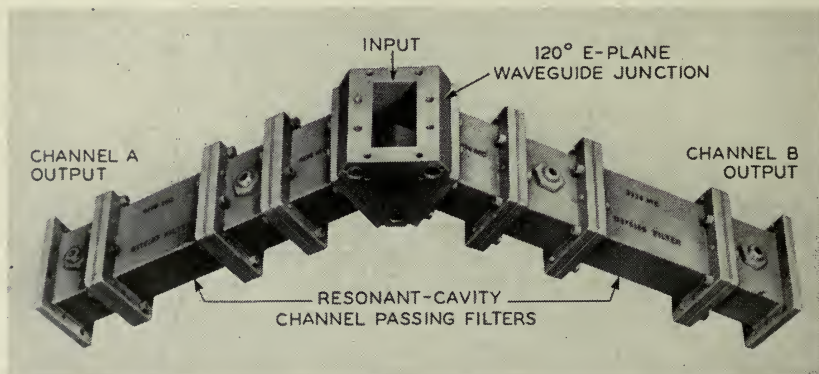


Fig. IV-2.—Photograph of a branching filter for an experimental radio relay system.

Several successful two-branch networks have been designed and constructed in this manner. One of these, developed for the New York-Boston circuit*, is illustrated in Fig. IV-2. Here two three-cavity filters are connected to an E plane Y junction, the waveguide analogue of a series connection. The two filters are tuned to different bands and each is connected to the junction through a line of length such that it causes no disturbance in the channel of the other. The electrical characteristics of this filter are plotted in Fig. IV-3.

Problems connected with the design of suitable microwave branching filters with more than two branches evidently differ considerably from previous filter problems. Channel passing networks which can be connected in series or parallel to form a channel branching filter can be designed at lower frequencies on the basis of lumped circuit theory and built of coils, condensers and resistances, but in the microwave region simple elements

* Developed by the group concerned with high-frequency filter design headed by A. R. D'heedene. A large part of the research underlying the design of these filters was performed by W. W. Mumford. Prior to the war a considerable amount of research on the band-pass type of waveguide filter had already been done by A. G. Fox.

and connections do not exist. Where more than two waveguide channel passing filters are to be connected to a common junction the design becomes complex, since in every channel the sum of the interactions of all the inactive filters on transmission through the active filter must be zero. It is evidently not easy to satisfy this condition, particularly since in doing so one must take account of the change with frequency of the effective length of all waveguide connecting lines. And even if such a solution is found it will be valid for only one set of channels, so that the problem must be solved all over again for every change in channel arrangement.

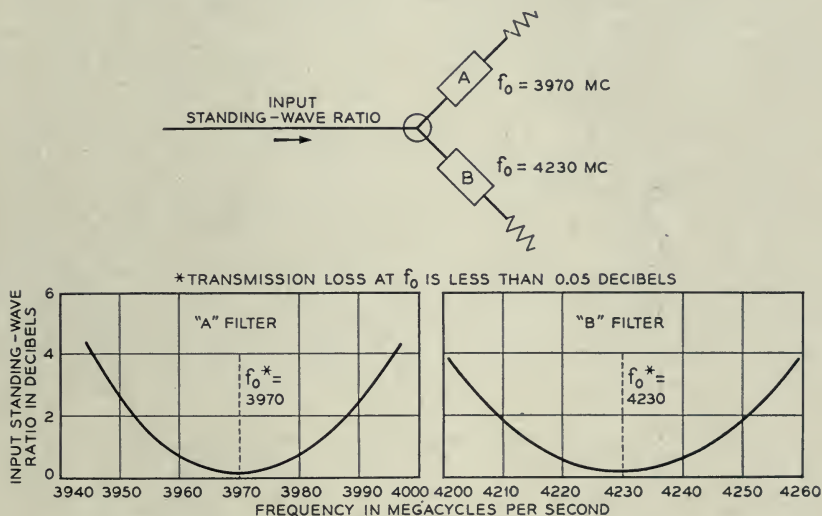


Fig. IV-3.—Input standing wave ratios of filter of Fig. IV-2.

As a result of these difficulties and after a few attempts to overcome them it became evident that a more flexible microwave branching filter technique should be found. Accordingly a solution on an iterative basis was developed. A channel dropping circuit was devised which, when inserted in a line, could extract or insert one channel while allowing others to pass through without disturbance. This circuit is of the constant resistance type; in other words it operates by diverting energy selectively but not by reflecting it back to the input. Consequently N such circuits placed in sequence do not interact reflectively; they, thus, form an N channel branching filter which is also of the constant resistance type.

An individual constant resistance channel dropping circuit is illustrated schematically in Fig. IV-4. It is made up of two hybrid²³ circuits, two

²³ For a general discussion of hybrid circuits see W. A. Tyrrell, "Hybrid Circuits for Microwaves", *Proc. I. R. E.*, vol. 35, No. 11, pp. 1294-1306, November 1947.

identical band reflection filters, and two quarter wavelengths of line. Each of the hybrids is analogous to a low-frequency hybrid coil and operates as follows. A wave in line C (See Fig. IV-4) incident on the hybrid is divided equally and with equal phase into A and B but does not appear in D or reappear in C. If waves in A and B are incident on the hybrid a wave proportional to their vector sum will appear in C, a wave proportional to their vector difference will appear in D but nothing will reappear in A or B. A wave in the input line incident on the channel dropping circuit will thus be divided by the input line into the lines leading to the two band reflection filters. These filters are designed to reflect frequencies lying within their band and pass all other frequencies. If the frequency is outside of the reflected band the two waves will travel on to connections A and B of the

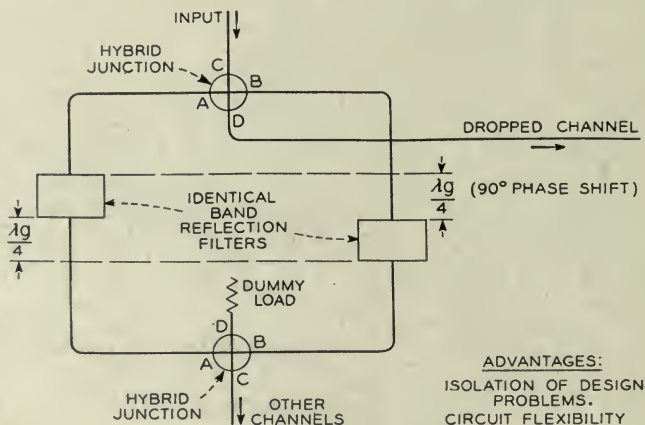


Fig. IV-4.—Schematic diagram of a constant impedance channel dropping filter using hybrid junctions and band reflection filters.

output hybrid. Here they will have equal phase and amplitude, their vector difference will be zero and the wave appearing in C of the output hybrid and consequently in the output line will contain all the power. If the frequency lies within the band of the reflection filters the two waves will be reflected by them and will travel back to the connections A and B of the input hybrid. The two waves strike these connections with opposite phase since one of them has traveled twice over an extra quarter wavelength of line. Their vector sum will consequently be zero and the wave which appears in terminal D of the input hybrid and consequently in the dropped channel line will contain all the power. The circuit of Fig. IV-4 is therefore a constant resistance channel dropping network which diverts energy lying within the band of the reflection filters but allows all other energy to pass through without disturbance. Conversely, by the law of reciprocity, this

circuit can insert energy lying within the band of the reflection filters into the main line without disturbing any energy passing through it at other frequencies.

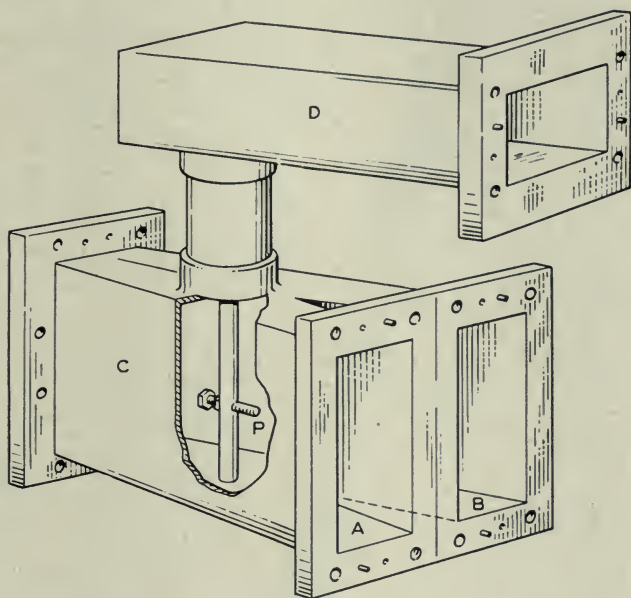


Fig. IV-5.—Hybrid junction used in the filter of Fig. IV-4.

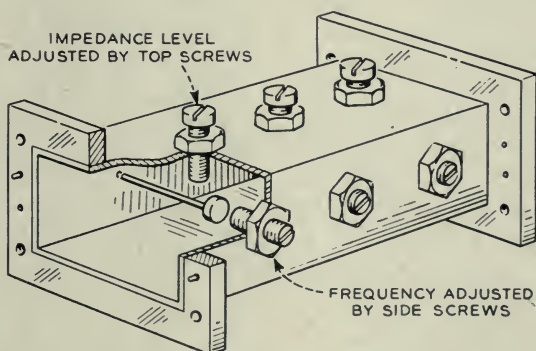


Fig. IV-6.—Waveguide band reflection filter used in the filter of Fig. IV-4.

An embodiment of the circuit of Fig. IV-4 suitable for use in the repeater arrangement of Fig. IV-1 has been constructed and tested. Figure IV-5 illustrates the waveguide hybrid employed. Here the waveguide opening

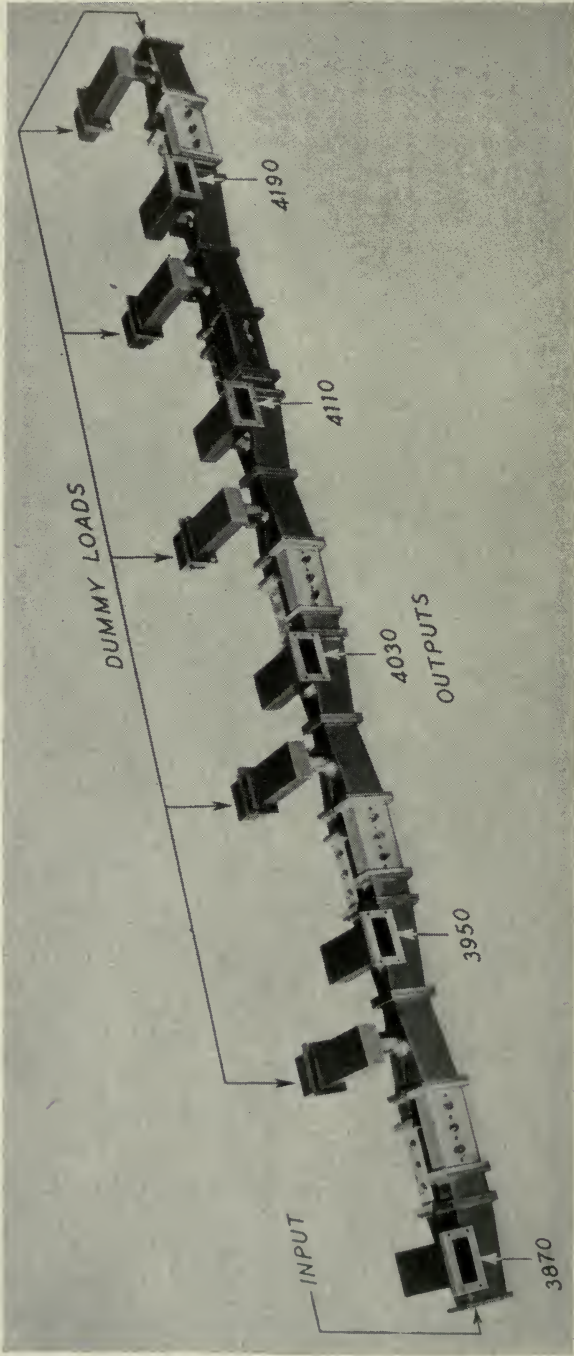


Fig. IV-7.—Photograph of a five channel branching filter.

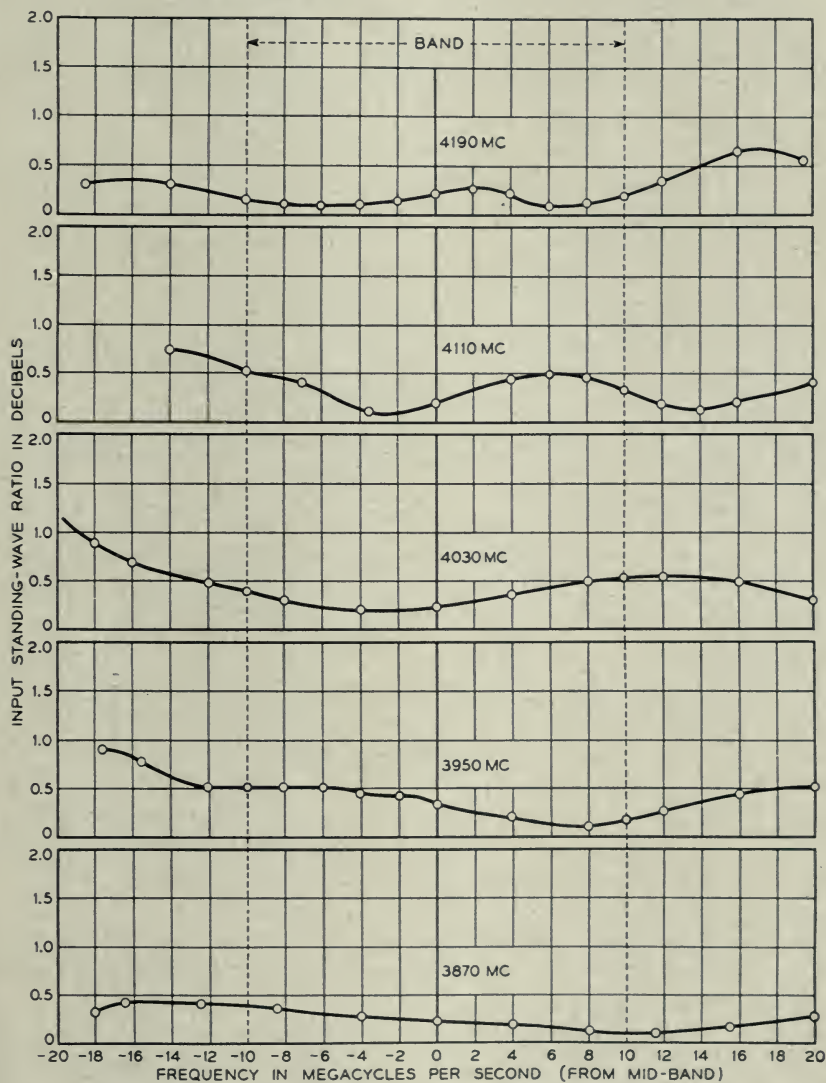
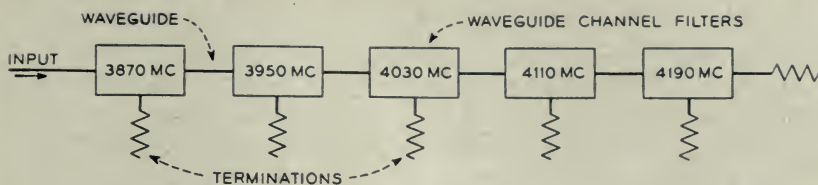


Fig. IV-8.—Input standing wave ratios for the filter of Fig. IV-7.

for C is physically parallel to those for A and B and is connected to them through a broad-band waveguide taper. Connection D is made through a relatively narrow band coaxial and probe arrangement. Figure IV-6 illustrates one of the waveguide band reflection filters. In this filter reflection occurs at three resonant rods, each tuned by an adjustable capacita-

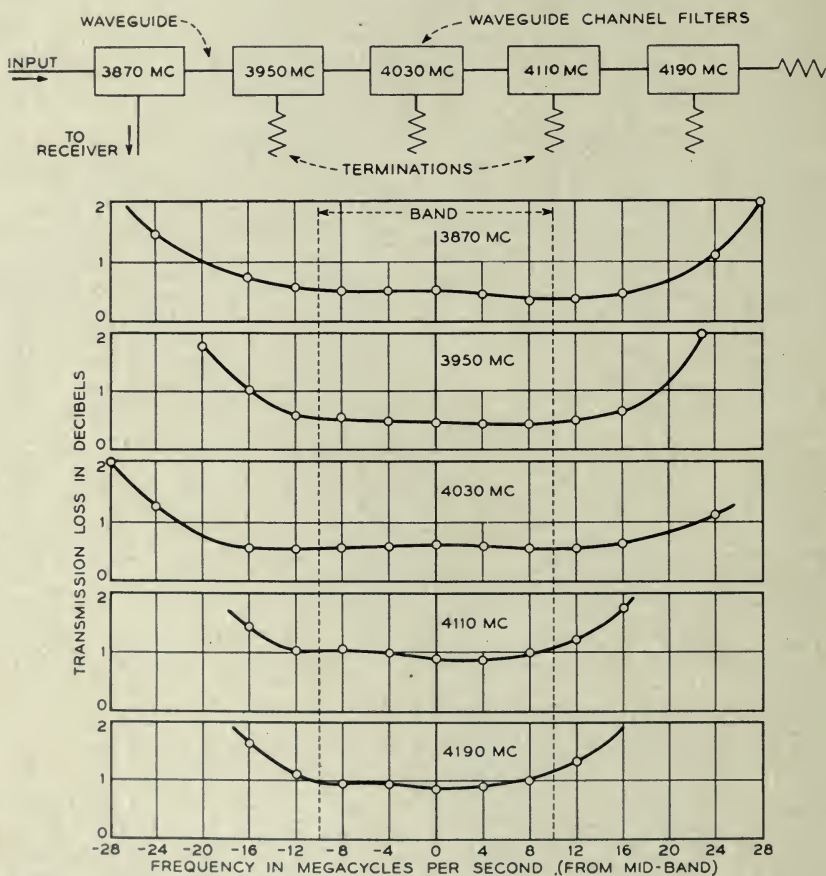


Fig. IV-9.—Transmission loss for the filter of Fig. IV-7.

tive plug at one end. These rods are placed perpendicular to the electric vector of the guide and are coupled to it by means of adjustable screws placed over them in the wide wall of the guide.

Channel dropping units for five channels in the 4,000 megacycle region were made up of these components and suitable quarter wavelengths of guide. These units were connected in sequence as shown in Fig. IV-7 and adjusted systematically. The electrical performance of the resulting five-

channel branching filter was measured with a double detection measuring set and is plotted in Figs. IV-8 and IV-9.

These measured electrical characteristics serve as a check on the general theoretical ideas concerning constant resistance hybrid branching networks. They also indicate that when these ideas are embodied in the form illustrated in Fig. IV-7 the result is a branching filter which can be used in currently planned radio relay circuits.

The circuit of Fig. IV-7 provides a satisfactory channel splitting network. It does not, however, provide consistent high off-frequency discrimination between one of the channel output terminals and the other terminals of the filter. When systems requirements* are such that extra discrimination or special impedance behavior is required, this can be supplied by inserting suitably designed reflecting filters in the branch lines. These added filters will not interact reflectively with the branching filter.

V. THE REPEATER AMPLIFIER†

In a microwave repeater circuit, Fig. II-1, the signal is amplified at each repeater to compensate for the transmission loss in the preceding radio path. Since we cannot build perfect amplifiers, the signal will not appear at the output terminals as a true replica of the input signal; the circuit will distort the shape of the signal and it will also add noise. Therefore, the main objectives in amplifier work have been to keep the signal distortion and the added noise within certain requirements.

To be more specific, the repeater amplifier in a relay system must be capable of supplying a maximum gain, G , as given by the equation II-1; it must have a ratio of output power capacity to noise figure which will meet the signal to noise ratio requirements of the system as given by equation II-4; since distortionless transmission is desired, it must have an amplitude characteristic as flat as possible and a phase characteristic as linear as possible over the essential range of frequencies of the signal it is desired to transmit;²⁴ and it must be equipped with an automatic gain regulating circuit to hold the output power constant over the expected range of input levels.

The simplest relay amplifier would be one which amplifies the signal and sends it on without a change in frequency. However, two major considerations indicated that early repeater amplifiers could not be so simple. No

* E.g., the converter may require a reflection in the input line at the image frequency, See Section V and Fig. V-4.

† Those parts of this section dealing with the general layout, the requirements, and the over-all testing of the repeater amplifier were prepared by D. H. Ring, who together with A. C. Beck did the work on this phase of the problem.

²⁴ Sallie Pero Mead, "Phase Distortion and Phase Distortion Correction", *B. S. T. J.*, vol. VII, No. 2, pp. 195-224, April 1928.

microwave amplifiers were known which gave promise of yielding an adequate ratio of output power capacity to noise figure, and there was considerable doubt of our ability to reduce, sufficiently, the feedback from the transmitting antenna to the receiving antenna.

These difficulties with straight-through amplification can be avoided by a repeater amplifier such as is shown schematically in Fig. V-1. The incoming signal is converted to an intermediate frequency, IF, where better amplifiers are available and where the major part of the required gain is supplied. The amplified IF signal is then converted back to the microwave frequency $f + \Delta f$, where Δf is relatively small. The difference Δf between the incoming and outgoing frequencies permits the use of circuit selectivity to counteract feedback troubles, and the radio frequency amplifier following the transmitting converter can have a relatively large noise figure.

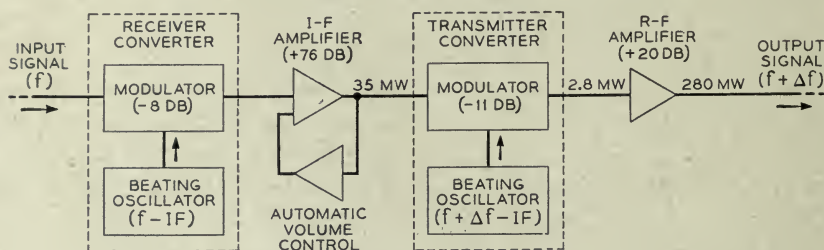


Fig. V-1.—Schematic of a repeater amplifier.

Our initial research on microwave repeaters was directed toward solving the problems associated with an amplifier of the type shown in Fig. V-1. The gain and level figures shown in this figure apply to the example of an eight-link relay system given in section II. They indicate approximate minimum objectives for the various components of the repeater amplifier to be discussed later.

Choice of I.F. Frequency—When selecting the intermediate frequency for a multichannel repeater circuit utilizing the interleaved radio frequency plan of Fig. IV-1 and the intermediate frequency type repeater amplifiers of Fig. V-1, the relative position of the various discrete frequencies and frequency bands shown in Fig. V-2 must be considered. In order to minimize the possibility of crosstalk from the image bands and interference from the beating oscillators caused by insufficient shielding, it is desirable to choose the intermediate frequency in such a way that the image bands fall midway between the active bands, and the oscillator frequencies fall midway between the image and active bands. These conditions are realized, as shown in Fig. V-2, if the intermediate frequency satisfies the relation

$$2 \text{ IF} = n\Delta f + \frac{\Delta f}{2}$$

or

$$\text{IF} = \frac{\Delta f}{4} (2n + 1)$$

where Δf is the frequency spacing and n is any integer greater than zero.

In general, better noise figures and circuit stability are obtained with low intermediate frequencies, while high intermediate frequencies lead to

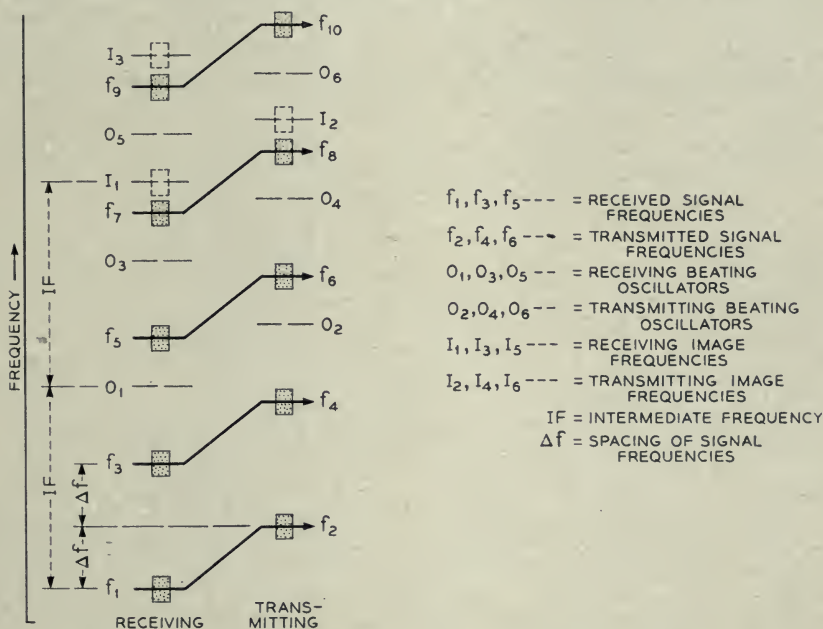


Fig. V-2.—Frequency spectrum of a multichannel microwave repeater circuit.

more symmetrical amplitude and phase characteristics. The research on the intermediate frequency components described below was conducted in the 60 to 70-megacycle range.

Frequency Stability—If all receiving and transmitting beating oscillators are independently controlled in a multichannel relay circuit of this kind, there is a possibility that small variations will add to produce large variations at the distant end of the system. This difficulty can be overcome by the frequency control system shown schematically in Fig. V-3. The transmitting and receiving beating frequencies are both derived from an oscillator

operating at a frequency suitable for receiving the incoming signal. The frequency of this oscillator is controlled by a high Q cavity and a servo mechanism to 0.2 megacycles or better.²⁵ One portion of the output of the oscillator is used as the beat frequency in the receiving converter. A second portion is combined with the output from a crystal oscillator operating at a frequency equal to the difference, Δf , between the incoming and outgoing frequencies. In this way a beat frequency for the transmitting converter is obtained which has the same variations as that for the receiving converter except for negligibly small variations that may occur in the crystal oscillator. As a result of this method of deriving the beat frequencies the outgoing frequency always differs from the incoming frequency by an amount equal to the crystal oscillator frequency and is not influenced by variations in the high-frequency local oscillator. The result is that, except for the small

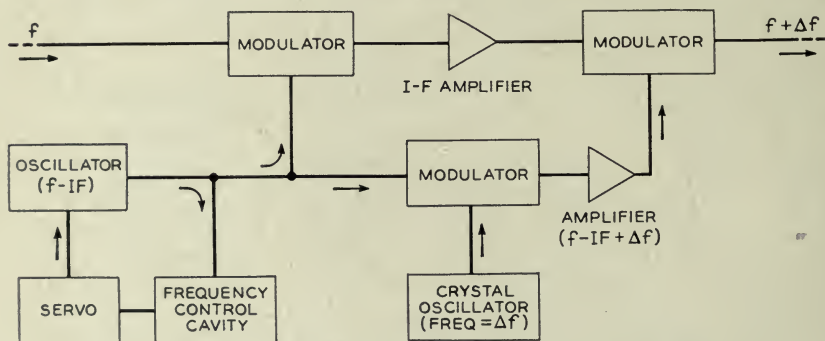


Fig. V-3.—Frequency control system for a microwave repeater.

crystal oscillator variations, all radiated frequencies in a long circuit carry only the variations in the transmitting oscillator of the originating terminal, while the intermediate frequency of each repeater may vary by an amount equal to the sum of the variation of its own local oscillator and that of the terminal transmitter frequency.

Automatic Gain Regulation—The function of the automatic gain control circuits is to hold the repeater output constant over the expected fading range. As already stated an allowance for fades 20 db down and 10 db up from free space have been made for 30-mile paths at a wavelength of about 7 centimeters. In addition to knowledge of the fading range to be compensated, it is necessary in the design of suitable circuits to know what the maximum fading rate is likely to be. Analysis of the records of a number of disturbed periods on the New York-Neshanic path indicate a maximum rate of 5 db per second.

²⁵ V. C. Rideout, "Automatic Frequency Control of Microwave Oscillators", *Proc. I. R. E.*, vol. 35, pp. 767-771, August 1947.

A conventional delayed automatic gain control circuit has been used in which a d.c. voltage, supplied by a peak rectifier in the output of the last IF stage, is amplified and fed back to bias several of the IF stages. Analysis of the transient response of the last repeater in response to a step function disturbance in the input to the first repeater, where the number of repeaters is of the order of 5 or more, shows that great care must be taken in shaping the frequency characteristic of the feedback circuit.

General Requirements of Components—The important factors bearing on the basic layout of the amplifier components have been discussed. Others, affecting the design of these various components will now be considered, after which a brief description of the research work on each component will be given.

Different repeater circuits will, in general, have different numbers of repeaters; also it may be necessary to feed signals into and extract signals from them at any point. Under these conditions it is impractical to specify only the over-all characteristics for a given number of repeaters. Each repeater itself should be individually good and should not depend upon any systematic compensation or equalization at any other point in the system. The repeater was designed in accordance with this philosophy and, in the interest of flexibility and ease of testing, the same line of reasoning was extended to cover the design of the components of the repeater as well.

In accordance with the above considerations major emphasis was placed on obtaining a minimum of amplitude variation and phase distortion over a 10-megacycle band for each repeater component. However, it was appreciated that even with the simplest circuits the inherent phase distortion would be excessive in long relay systems. Phase equalizers can be designed to equalize this distortion, but the difficulties of design and alignment increase with the magnitude of the distortion to be equalized. Accordingly, simple circuits were used wherever gain requirements would permit and at the same time parallel research was carried out on the problems of designing and testing appropriate phase equalizers.

While our aim was to provide a repeater 10 megacycles wide suitable for any type of modulation, it soon became apparent that it would be uneconomical to attempt to provide the extreme degree of linearity that would be required; for example, for a long relay circuit carrying an amplitude-modulated television signal. However, early tests indicated that very satisfactory transmission could be obtained with low-index frequency-modulated signals, for which reason the later stages of the work were aimed at providing an amplifier to be used for such signals. Nevertheless an attempt was made to limit the compression in each unit except the R.F. amplifier to 0.1 db at maximum rated load.

A further important requirement placed on all components was that their input and output impedances should match the corresponding impedances of the components to which they were to be connected with a minimum of reflected power over the 10-megacycle band. This requirement was imposed on the separate units to provide for flexibility in testing and to permit easy patching in of spare units in case of failure.

RECEIVING CONVERTER*

The receiving converter, together with the input to the first stage of the intermediate-frequency amplifier, occupies a unique position in the repeater amplifier in that it is located at that point in the circuit where the signal level is lowest. As a consequence, research on receiving converters has been directed toward insuring that the least possible noise be added to the signal to be amplified. An extensive background of microwave converter design information was available from the work done on converters during the war^{26,27} which led to the selection of a balanced converter using a waveguide hybrid junction and 1N23-B silicon point contact rectifiers. The information already available would probably have been adequate except for the two additional requirements imposed by the repeater amplifier: first, that the standing wave ratio at the input have a low value and, second, that uniform conversion efficiency be maintained over a band of at least 10 megacycles.

A receiving converter with its associated input and output circuits is shown in Fig. V-4. Frequency conversion is accomplished in a device of this kind by virtue of the fact that when two sinusoidal voltages (in this case the signal and the beating oscillator) are applied to a non-linear impedance such as a rectifier, new frequencies given by the sums and differences of the applied frequencies and their harmonics are generated. The difference frequency may thus be selected as the desired output frequency and passed through the IF transformers to the IF amplifier. The performance of a converter is, however, influenced by the impedances encountered by some of the other frequencies generated. For this reason, the separation *S* between the input filter, more properly a component of the channel selecting network, but here shown as part of the converter, and the converter must be given consideration, since it determines the phase of the reflection from the filter at the image frequency (the difference between the beating

* This section prepared by C. F. Edwards who was responsible for the research on the receiving converters.

²⁶ "Developments of Silicon Crystal Rectifiers for Microwave Radar Receivers", J. H. Scaff and R. S. Ohl, *B. S. T. J.*, vol. 26, pp. 1-30, January 1947.

²⁷ Descriptions of several converters developed prior to and during the war as well as a more complete description of the present converter are given in C. F. Edwards' paper, "Microwave Converters", *Proc. I. R. E.*, vol. 35, No. 11, pp. 1181-1191, November 1947.

oscillator second harmonic and the signal) which in turn affects the converter output impedance and hence the match between the rectifiers and the IF transformers. Failure to obtain a proper match results in non-uniform transmission over the channel band.

In addition, to maintain uniform conversion efficiency over a wide band, it is necessary to control the impedance encountered by frequencies near the beating oscillator second harmonic. This is done by means of the harmonic

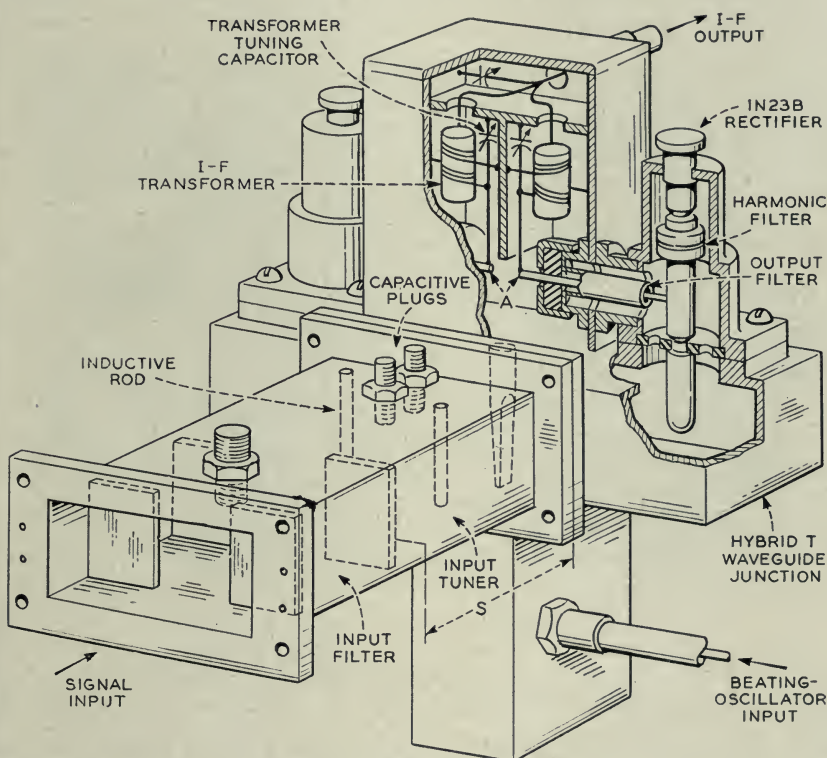


Fig. V-4.—Receiving converter with input filter.

filters shown which reflect these frequencies from a point close to the rectifier. Since the point of reflection is close, there is little opportunity for the phase of the reflection, and hence the conversion loss, to vary over the band of frequencies to be converted.

The converter in Fig. V-4 was designed to provide as good a termination as possible to the incoming waveguide over the band of channel frequencies so that a minimum of additional adjustment would be required to match the guide accurately at any particular channel frequency. This additional

adjustment is provided by the input tuner shown. With it, the reflection coefficient may be reduced to zero at the desired channel band center, and when this is done the standing wave ratio at the input is less than 2 db over a 20-megacycle band.

The bandwidth of the converter is largely determined by the IF transformers shown which transform the balanced output impedance to the unbalanced 75-ohm coaxial line connecting to the IF amplifier. When transformers having a coupling coefficient of about 0.5 are used, transmission variations of less than 0.1 db over a 20-megacycle band are obtained.

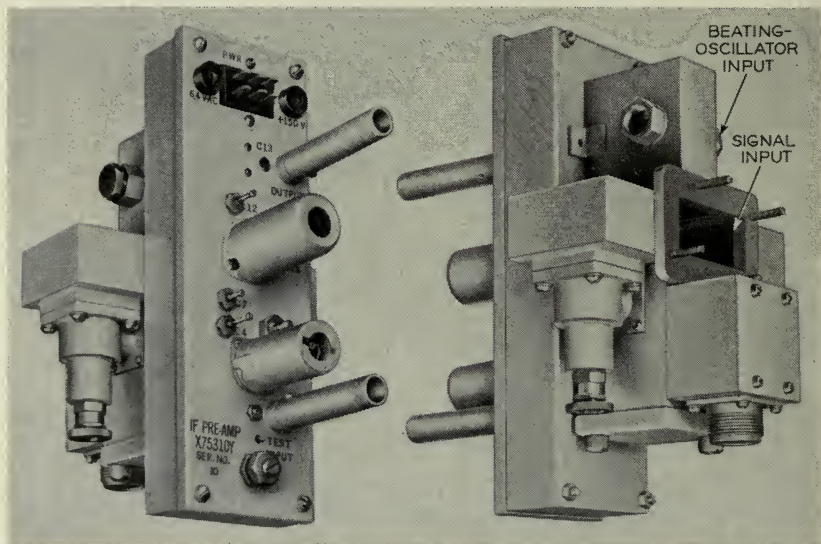


Fig. V-5.—Receiving converter with low noise IF preamplifier.

The noise figure of the repeater amplifier is determined by the conversion loss and noise figure of the converter and the noise figure of the IF amplifier.¹⁶ The converter designed for the New York-Boston circuit has a conversion loss of about $5\frac{3}{4}$ db and its noise figure is 10 db. Thus when it is used with an IF amplifier having a noise figure of 7 db, a figure of 14 db is obtained for the over-all noise figure of the repeater amplifier. Figure V-5 shows this converter with a low noise preamplifier attached.*

I.F. AMPLIFIER†

Most of the gain of the IF amplifier is obtained with stages using double tuned, symmetrically loaded (or 'matched') interstage transformers, de-

¹⁶ Loc. cit.

* Developed by H. C. Foreman and B. C. Bellows, Jr.

† This section prepared by Karl G. Jansky who, together with V. C. Rideout, did the work on IF amplifiers.

signed in accordance with the formulas given by a former member of this laboratory.²⁸ In Fig. V-6, "a" shows a schematic diagram of such a transformer and "b" shows the equivalent π network generally used. The equivalent T network shown at "c" has also been used. In Fig. V-7 "a" shows the theoretical band-pass characteristic for this type of transformer with a coupling coefficient of 0.5 which was the value used in most stages. The circles indicate points measured on a typical transformer. This matched network is relatively insensitive to small changes in capacitance so that wherever it is used it is possible to change tubes without having to realign the amplifier. The delay distortion for this type of network is, as shown by

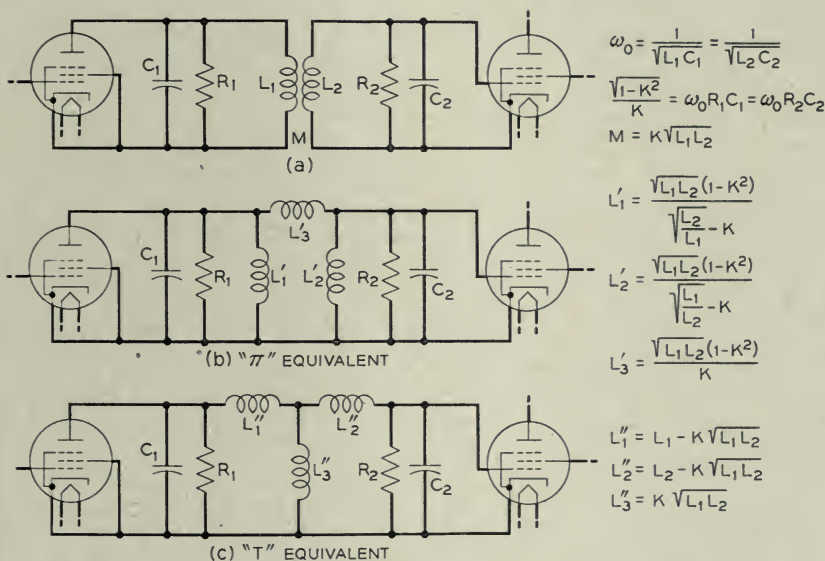


Fig. V-6.—The double tuned IF transformer.

"b" in this figure, also relatively small. The gain per stage for a coupling coefficient of 0.5 is approximately 6 db for 6AK5 vacuum tubes and about 12 db for the recently developed WE 404-A tubes. As shown by W. J. Albersheim²⁸ it is possible to design circuits which will give more gain than the "matched" transformer, but only at the cost of increased sensitivity to capacitance changes. For convenience the IF amplifier is usually divided into a preamplifier closely associated with the receiving converter and a main IF amplifier.

Pre-amplifier—At the time work was begun, noise figures of the order of

²⁸ V. C. Rideout, "Design of Parallel Tuned Transformers", *B. S. T. J.*, vol. 27, pp. 96-108, January 1948.

12 to 14 db were obtainable for 65 mc amplifiers of the required bandwidth with 6AK5 tubes and matched input transformers. This was much worse than was desired. By using a 6J4 close spaced grounded grid triode in the input stage, much lower noise figures were obtained, but the gain with matched input and output circuits was so low (approximately 3 db for a coupling coefficient of 0.5) that the noise from the following stage contributed considerably to the overall noise figure.

By removing the loading resistance on the output side of the first inter-stage transformer and reducing the coupling coefficient to 0.3, the gain was

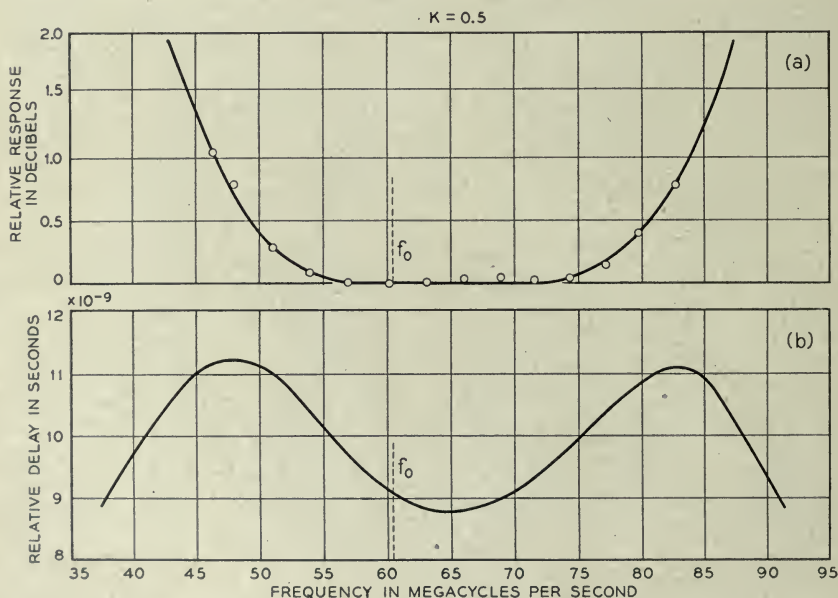


Fig. V-7.—Band pass and delay characteristics of the IF transformer.

raised sufficiently to give an overall noise figure of the order of 7 db. Figure V-5 shows the preamplifier developed for the New York-Boston circuit. This amplifier employs a 6J4 and a WE 404-A and provides a gain of 23 db.

The 6J4 tube, when used in a grounded grid circuit, has an input impedance of approximately 85 ohms which is close to the desired 75 ohm impedance. When a good match is required, it is necessary to use an input transformer, but it should be noted that an improvement in the noise figure can be obtained by deliberately producing a mismatch in the right direction at the input. Noise figures as low as 4 db have been obtained in this manner with recent experimental tubes.

Figure V-8 is a schematic diagram of an amplifier with a very low noise

figure consisting of two of these experimental grounded grid tubes in tandem. The circuits were designed so that there is a mismatch at the input of each tube. The overall noise figure is 4.0 db with the amplifier connected between a 75-ohm generator and a 75-ohm load; the gain is $17\frac{3}{4}$ db; and the bandwidth at the 0.1 db down points is about 13.5 megacycles.

Main IF Amplifier—In order to obtain the required output power over the desired bandwidth it was necessary, at the beginning of this work, to use the WE 367-A tube in the output stage. Since the gain of a stage using this tube is very low, it must be driven by a tube similar to the 6AG7 and to prevent compression in this latter tube a special high gain, triple-tuned interstage transformer was designed, with a relatively low coupling coefficient.

Methods of paralleling tubes to get more power output or the same power output with a much wider bandwidth have been worked out, but recent

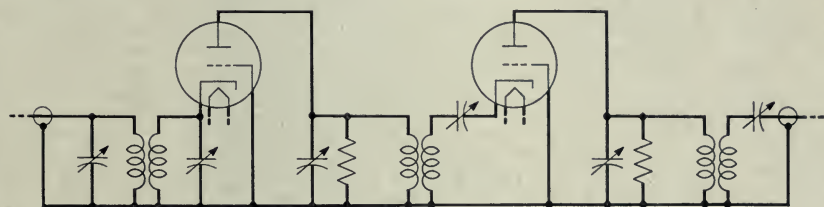


Fig. V-8.—Schematic of a "low noise figure" preamplifier.

developments in power output tubes and in transmitting converters have made them unnecessary, except for applications requiring very wide bandwidths.

Automatic gain control can be applied to the IF amplifiers by controlling the grid bias of the various stages. However, it is not advisable to apply the gain control bias to either of those stages which are preceded by the special high-gain transformers. In these stages the slight changes in input impedance of the tubes caused by variations in the grid bias would be sufficient to alter significantly the band-pass characteristics of the transformer.

Figure V-9 shows an amplifier with about 55 db gain developed for the New York-Boston circuit.* Automatic gain control bias is applied to the grids of the first three stages which employ wide band ($K = 0.7$) matched interstage transformers.

It will be noted that the first interstage transformer of the preamplifier and the last interstage transformer of the main IF amplifier require low coupling coefficients to obtain the gain desired at these points. For this

* Developed by A. L. Hopper and B. C. Bellows, Jr.

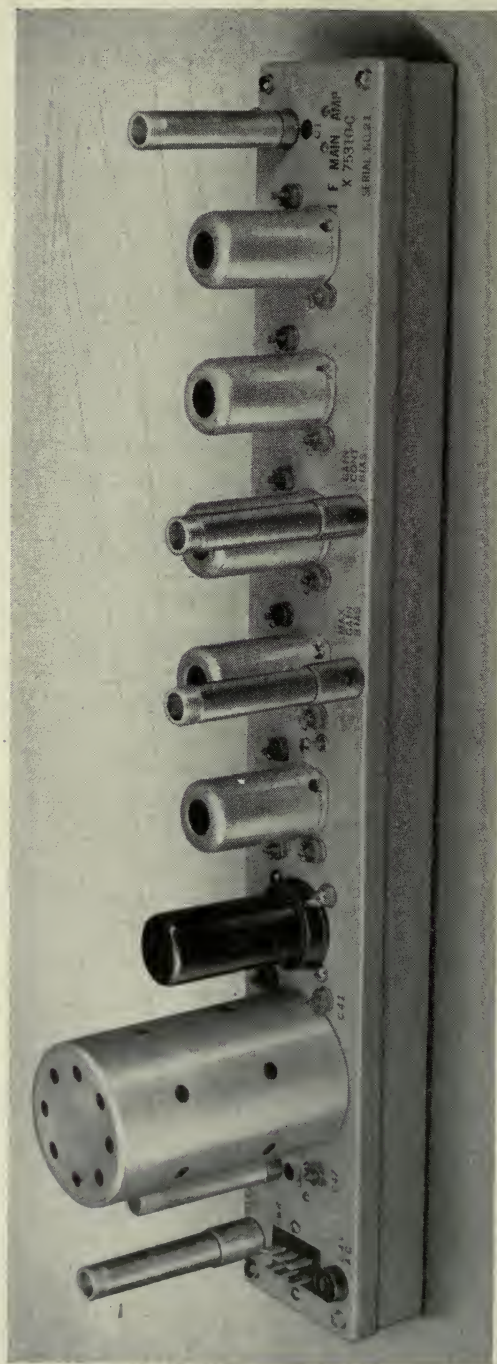


Fig. V-9.—The main If amplifier.

reason, these two transformers largely determine the band pass and delay distortion of the whole amplifier. Typical overall characteristics for a complete IF amplifier are shown in Fig. V-10.

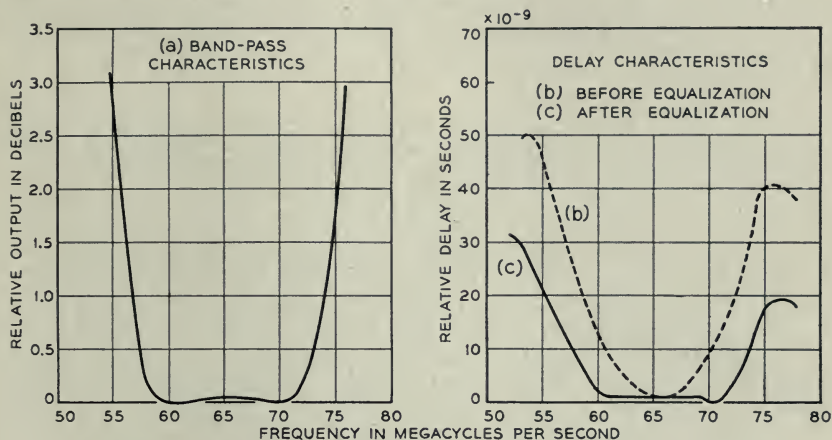


Fig. V-10.—Band pass and delay characteristics of complete IF amplifier.

TRANSMITTING CONVERTER, OR MODULATOR*

The transmitting converter problem differed from the first converter problem in that high output power was the major consideration, rather than low conversion loss and noise figure.

There will normally be three frequencies present in the output of a converter. These are the desired output sideband frequency, f' , the beating oscillator frequency, $f' + \text{IF}$ or $f' - \text{IF}$, and the unwanted sideband frequency, $f' + 2 \text{IF}$ or $f' - 2 \text{IF}$. The strongest of these is the beating oscillator frequency, but fortunately this can be suppressed by 20 to 30 db by balance in a balanced converter. Investigation showed that the input impedance at the IF terminals of the modulator was affected by the load impedance of the modulator at both the wanted and the unwanted sideband frequencies. The load impedance consists of the input impedance of the RF amplifier as seen through the length of transmission line which connects the two components. At the wanted sideband frequency, the RF amplifier presents a good termination and the load impedance is independent of the length of the connecting line. At the unwanted sideband frequency, however, the RF amplifier presents a short circuit and hence the load impedance is a function of the length of the connecting line. A waveguide filter was incorporated in the output circuit of the modulator so as to reflect the un-

* This section prepared by W. W. Mumford who did the work on the transmitting converters.

wanted sideband back into the modulator in proper phase, thereby making the length of the line going to the RF amplifier much less critical. The proper phase of reflection was obtained by adjusting the length of the waveguide between the modulator and the reflecting filter. The IF impedance of the crystals could thus be varied so as to obtain an impedance match between the IF amplifier and the crystals.

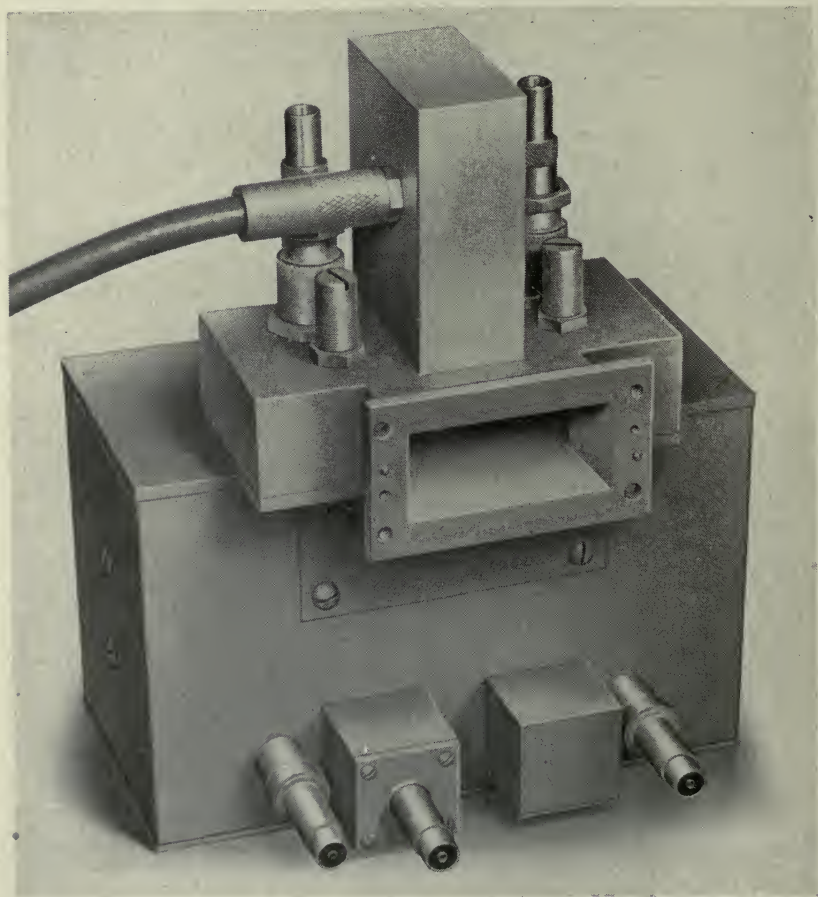


Fig. V-11.—Experimental model of transmitting converter.

An investigation of the power capacity of various types of standard and special silicon crystals indicated that one similar to the 1N28 crystal was the best available for this service, and life tests were made to check the stability of the crystal under high level conditions. Special IF input and crystal-to-waveguide matching circuits were developed to meet the stringent

match requirements. The research on these major factors and many subsidiary details resulted in a converter which had an output of 6 milliwatts with less than 0.1 db compression, good input and output match, a flat amplitude response over more than 10 megacycles, and a conversion loss of about 11 db.

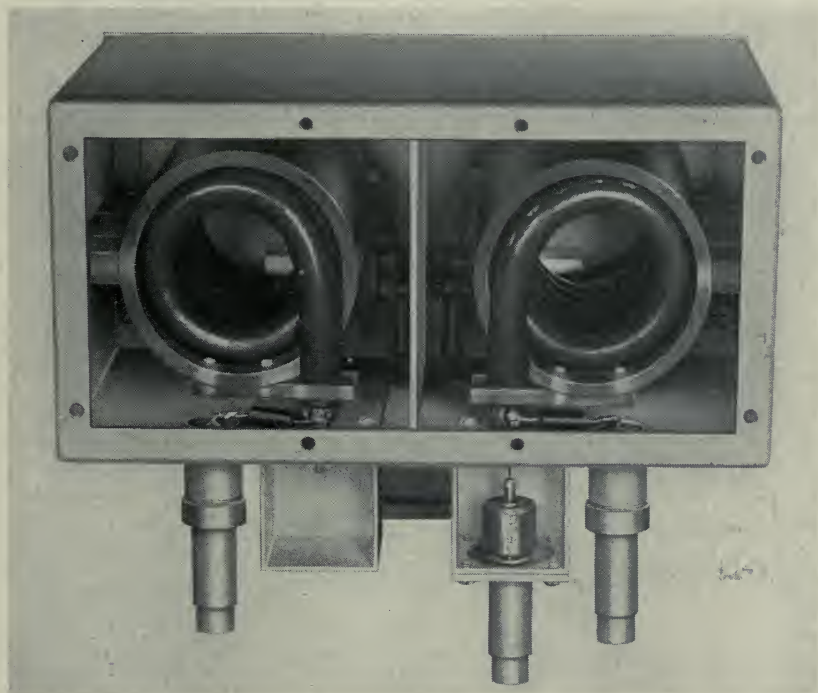


Fig. V-12.—Unbalanced to balanced coaxial transformer for feeding the 65 mc signal to the transmitting converter.

This converter, or transmitting modulator, had the 1N28 crystals mounted in the conjugate branches of a waveguide hybrid junction, as shown in Fig. V-11. Adjustable coaxial sleeves surrounded the crystal cartridges in order to effect an impedance match to the waveguide, and tuning studs were provided for trimming adjustments. The beating oscillator was injected into the hybrid junction through a broad-band coaxial-to-waveguide transducer in the upper branch of the hybrid junction, and the sidebands appeared in the conjugate waveguide branch. The 65-megacycle signal was fed onto the crystals in push-pull through the unbalanced to balanced coaxial transformer shown in Fig. V-12. Blocking condensers enabled the crystal currents to be monitored separately and RF filters kept the 4000-megacycle

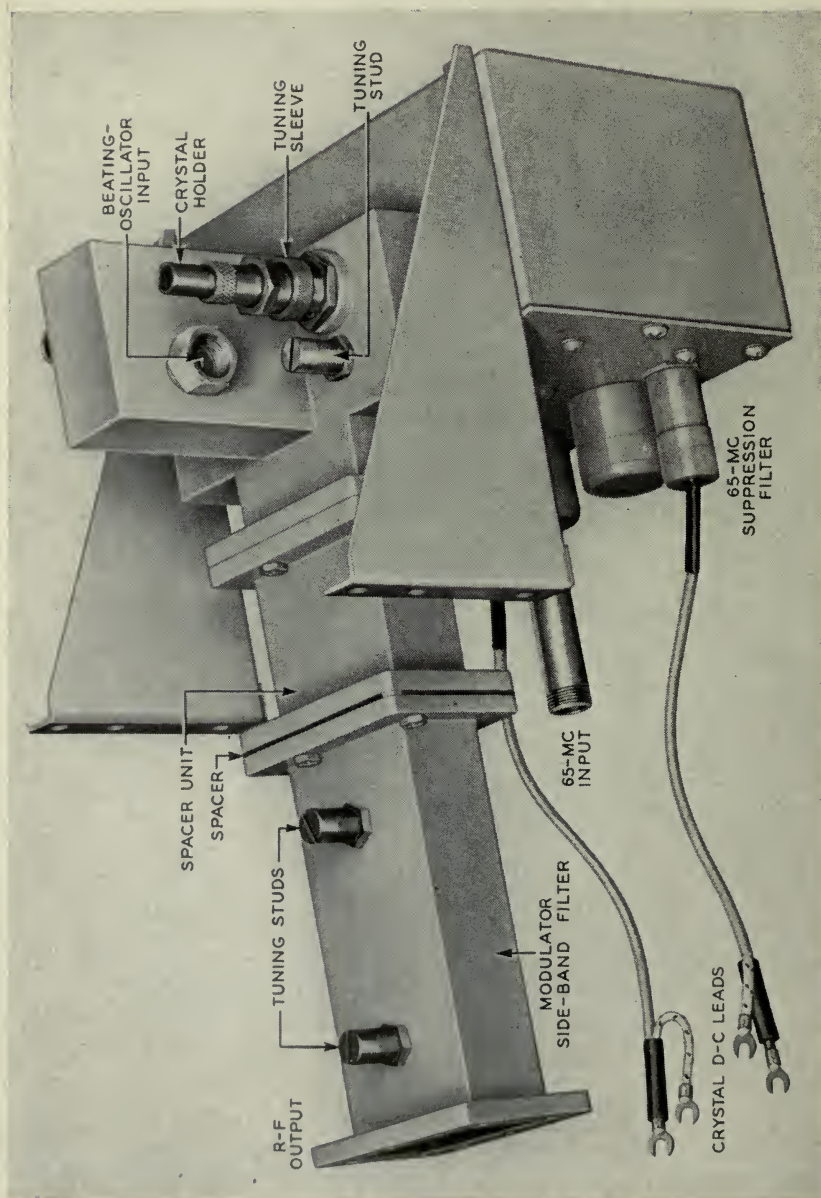


Fig. V-13.—Transmitting converter complete with output filter.

energy from entering the 65-megacycle transformer compartment. The converter developed for the New York-Boston circuit, complete with sideband filter, spacer, mounting brackets and d-c. leads, is shown in Fig. V-13.*

R. F. AMPLIFIER**

Prior to the war, a considerable amount of research had been applied in the Bell Telephone Laboratories to electron tubes operating on the velocity modulation principle, with the expectation that such tubes would find applications in radio-relay systems. Although this work was interrupted by the war, enough information had been obtained to make it apparent, upon resumption of the radio relay work, that such tubes were the only ones then known which showed promise of meeting the stated requirements.

Velocity modulation tubes have been described by several authors,^{29, 30, 31} and the theory of their operation has been discussed adequately in several places.^{32, 33, 34, 35} However, a review in 1939 of the structures then known had led to the decision that a new type of construction would be necessary to obtain a satisfactory amplifier tube for radio relay purposes. To keep the tube voltages within reasonable limits it is desirable to make the input and output gaps as small as possible. Resonant circuits external to the evacuated envelope are desirable to enable coverage of as large a frequency range as possible with a single tube, and to facilitate addition of broad-banding circuits. Grids on the input and output gaps are undesirable because of the large interception of current and the difficulty and expense of construction.

With the above considerations in mind, decision was then made to explore the possibilities of focussed beams, and of gaps comprised of copper discs sealed through a cylindrical glass vacuum envelope. The latter technique[†] had been developed at the Bell Telephone Laboratories in connection with

* Developed by H. C. Foreman and W. W. Halbrook.

** This section prepared by A. G. Fox and A. E. Bowen. Messrs. Fox and Bowen, in collaboration with A. L. Samuel, A. E. Anderson, and J. W. Clark of these Laboratories, did the major part of the research which resulted in this amplifier.

²⁹ Varian, R. H. and Varian, S. F., "A High Frequency Oscillator and Amplifier", *Jour. Applied Physics*, vol. 10, No. 5, pp. 321-327, May 1939.

³⁰ Hahn, W. C. and Metcalf, G. F., "Velocity Modulated Tubes", *Proc. I. R. E.*, vol. 27, No. 2, pp. 106-116, Feb. 1939.

³¹ Harrison, A. E., "Klystron Tubes", McGraw-Hill Book Co., New York, 1947. (Book)

³² Hansen, W. W. and Richtmyer, R. D., "On Resonators Suitable for Klystron Oscillators", *Jour. Applied Physics*, vol. 10, No. 3, pp. 189-199, March 1939.

³³ Hahn, W. C., "Small Signal Theory of Velocity Modulated Electron Beams", *G. E. Review*, vol. 42, No. 6, pp. 258-270, June 1939.

³⁴ Hahn, W. C., "Wave Energy and Transconductance of Velocity-Modulated Electron Beams", *G. E. Review*, vol. 42, No. 11, pp. 497-502, Nov. 1939.

³⁵ Ramo, S., "The Electronic-Wave Theory of Velocity Modulation Tubes", *Proc. I. R. E.*, vol. 27, No. 12, pp. 757-763, Dec. 1939.

the development of water cooled tubes, and showed promise of providing a very convenient means for construction of resonant circuits in which a part of the circuit was external to the vacuum envelope and part was internal.

Numerous forms of focussed beam-disc seal velocity modulation tubes were constructed for examination of the various factors affecting the performance of such tubes. Sizes and shapes of gaps, length of drift space, voltage and beam current, and other factors were adjusted to optimize the performance. Both magnetic and electrostatic focussing of the beam were explored. Triple gap tubes, from which gains of more than 30 db were obtained, were experimented with.

The end result of this work was the development of a four-stage amplifier employing velocity modulation tubes of the disc-seal type (Fig. V-14) designed especially for this application.* The electron gun is at the left, and the collector is at the right. These tubes employ external circuits in the form of resonant cavities assembled around the tube. The electron

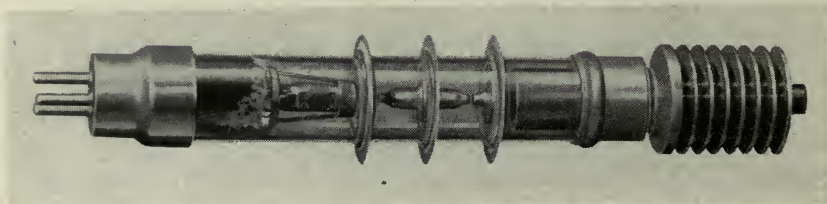


Fig. V-14.—Disc seal velocity modulation tube.

beams are accelerated by 1500 volts and focussed magnetically by means of small alnico permanent magnets placed around the cavity structure as shown in Fig. V-15. This shows a single stage of coaxial-coupled amplifier with half of each cavity removed to show internal details. The cavities are tunable over approximately a 250-megacycle range by means of metal screw plugs located around their periphery. Two different sets of cavities are sufficient to cover the entire frequency range of the radio repeater.

A single amplifier stage of the type shown in Fig. V-15 will exhibit a single-tuned type of gain characteristic which is only 5 megacycles wide at the 3 db points when operated under matched input and output impedance conditions. It was therefore necessary to widen the band by using double-tuned circuits throughout. In fact such a design would be difficult to avoid because each tube has its own resonant cavities. By running a short length of waveguide or coaxial transmission line from the output cavity of one stage to the input cavity of the next, a double-tuned structure results automatically. The coupling from cavity to transmission line is made by means

* These tubes were developed by A. L. Samuel and J. W. Clark.

of a window in the side of the cavity when a waveguide line is used, and by means of a wire loop projecting within the cavity when a coaxial line is used. Amplifiers employing both types of coupling were designed and tested. The coaxial coupling type with links approximately one-half wavelength long

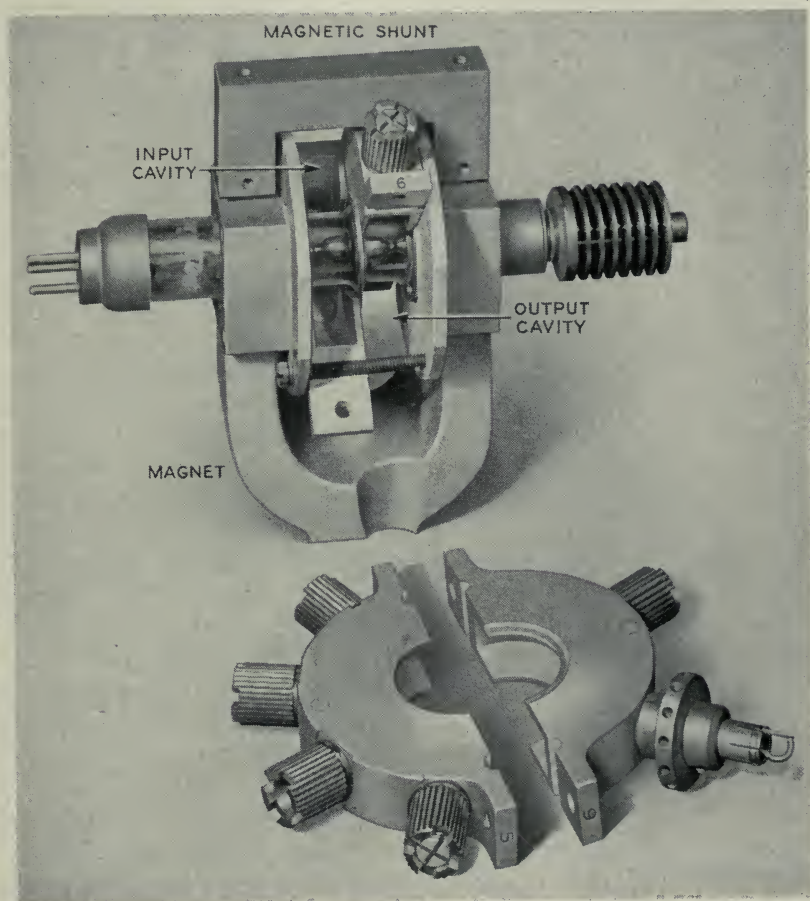


Fig. V-15.—A single stage of a coaxial-coupled amplifier using the velocity modulated tube of Fig. V-14.

proved the easier to adjust. Variation of the coupling from zero to a maximum is accomplished by rotating the plane of the coupling loop through 90° . Each cavity is provided with a small plate of resistance film projecting into the field in such a way that rotation of the plate about an axis lying in the plane of the plate will vary the resistive loading present in the cavity from zero to a maximum. Finally, the input and output cavities of

the first and last stages respectively are coupled to separate tuned cavities in order to provide double-tuned terminals for the amplifier. This results in the overall structure shown schematically in Fig. V-16.

The required bandwidth is obtained by a process of stagger-coupling the circuits so that the individual responses are as indicated schematically at the top of Fig. V-16. Because there are available continuously variable adjustments of tuning, loading, and coupling for each circuit, it is possible to obtain a very smooth and symmetrical overall response characteristic as shown in Fig. V-17. The corresponding measured delay distortion characteristic is shown in Fig. V-18. For this type of tuning the output power is about .7 watt at $\frac{1}{2}$ db of compression. More power is, however, obtainable if more compression is tolerable; and when used in the repeater for the transmission of FM signals, the amplifier is driven to an output of 1 watt.

Since the circuits are of fairly high Q , the frequency characteristics of the amplifier are markedly affected by changes in temperature of the cavities. In order to minimize such detuning effects, the amplifier has been placed in a temperature controlled compartment. Since about 45 watts are dissipated at the collector of each tube by the high-voltage electron beam, the collectors must be cooled by a forced air blast. In order to keep the cooling system separate from the temperature control system, the collector ends of the tubes project through a wall of the temperature controlled compartment into an external air duct. Figure V-19 shows a complete r.f. amplifier with the cover of the temperature control box removed. This is the amplifier developed for the New York-Boston circuit.* The electron-gun ends of the tubes face the reader. Short lengths of flexible coaxial cable couple the input and output of the amplifier to the associated equipment.

Testing of Components and Repeater Amplifier—Many special measuring and testing techniques had to be devised for the research work on each component. In addition, standard production tests had to be worked out and measuring equipment constructed.

Impedances were measured at RF with standing wave detectors, and at IF with three fixed taps on a coaxial line. It was found that input and output SW ratios could be held to 1.7 db or less over the 10 megacycle band. Both point by point and swept oscillator methods were used at RF and IF for measuring amplitude characteristics. Particularly as a result of the development of swept oscillator measuring techniques it was found practicable to adjust each unit to ± 0.1 db amplitude variation over the 10 megacycle band. Noise figures¹⁶ of IF equipment were measured by the noise diode method.³⁶

* Developed by F. E. Radcliffe and R. C. Carlton.

¹⁶ Loc. cit.

³⁶ H. Johnson, "A Coaxial Line Diode Noise Source for U.H.F.," *R. C. A. Review*, vol. VIII, No. 1, pp. 169-185, March 1947.

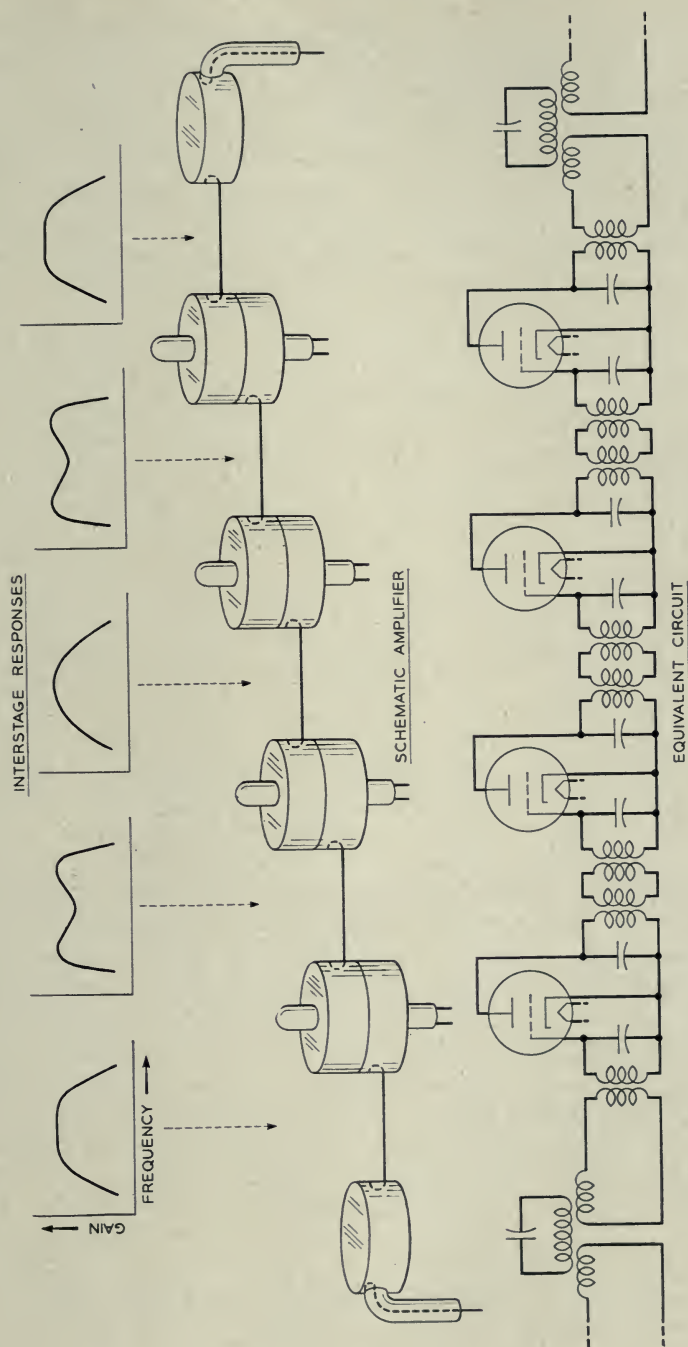


Fig. V-16.—Schematic diagram of four stage coaxial coupled amplifier using velocity modulated tubes.

A method³⁷ of measuring the phase and group delay characteristics of the repeater was worked out. This is a rather difficult measurement to make because the tolerable phase and delay distortion are very small due to the wide band of the system. Relative group delay through the band was measured with an accuracy of about ± 0.001 microsecond. This corresponds to an accuracy in relative phase shift of about $\pm 0.35^\circ$. The measured distortion agreed reasonably well with the distortion calculated from the constants of the various circuits. These measurements and calculations showed that while the characteristics of an 8-link repeater

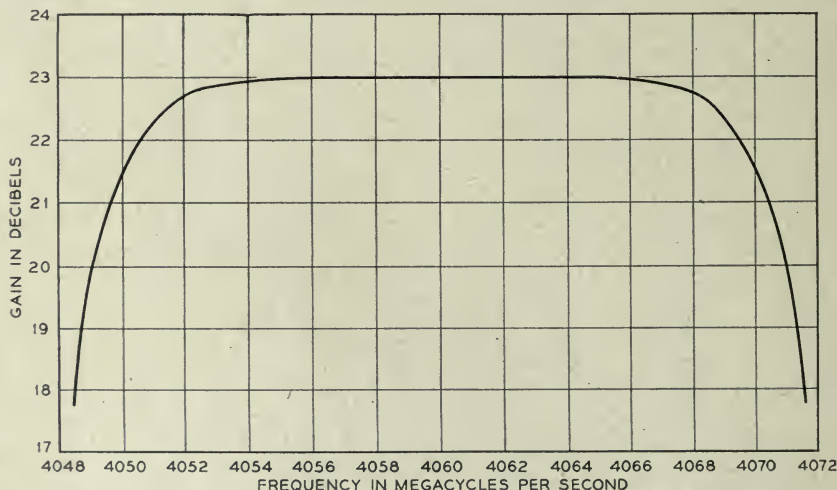


Fig. Y-17.—Frequency response of RF amplifier.

circuit will give acceptable television pictures, phase equalization will provide a definite improvement. Longer circuits will therefore require phase equalization. Equalizers were devised for both the IF and RF circuits which made it possible to equalize the group delay of each component of the repeater amplifier to ± 0.001 microsecond over the 10-mc band.

An experimental repeater amplifier was set up so that the output could be fed through an attenuator to the input. It was then possible to break the loop and make overall tests of delay, amplitude linearity, and transient response with IF measuring equipment. Transient response was measured by including a 2000-foot waveguide line between the input and output terminals and applying the circulated pulse testing technique.³⁸ This

³⁷ D. H. Ring, "The Measurement of Delay Distortion in Microwave Repeaters", elsewhere in this issue of the *B. S. T. J.*

³⁸ A. C. Beck and D. H. Ring, "Testing Repeaters with Circulated Pulses", *Proc. I. R. E.*, vol. 35, No. 11, pp. 1226-1230, November 1947.

testing method permits a study of the shapes of rectangular pulses which have passed through a repeater amplifier many times, and thus simulates transmission through a relay system with many repeaters. Figure V-20 shows an example of the results obtained from such tests made on the IF amplifier alone. The top row shows a 1.25-microsecond test pulse after 1,

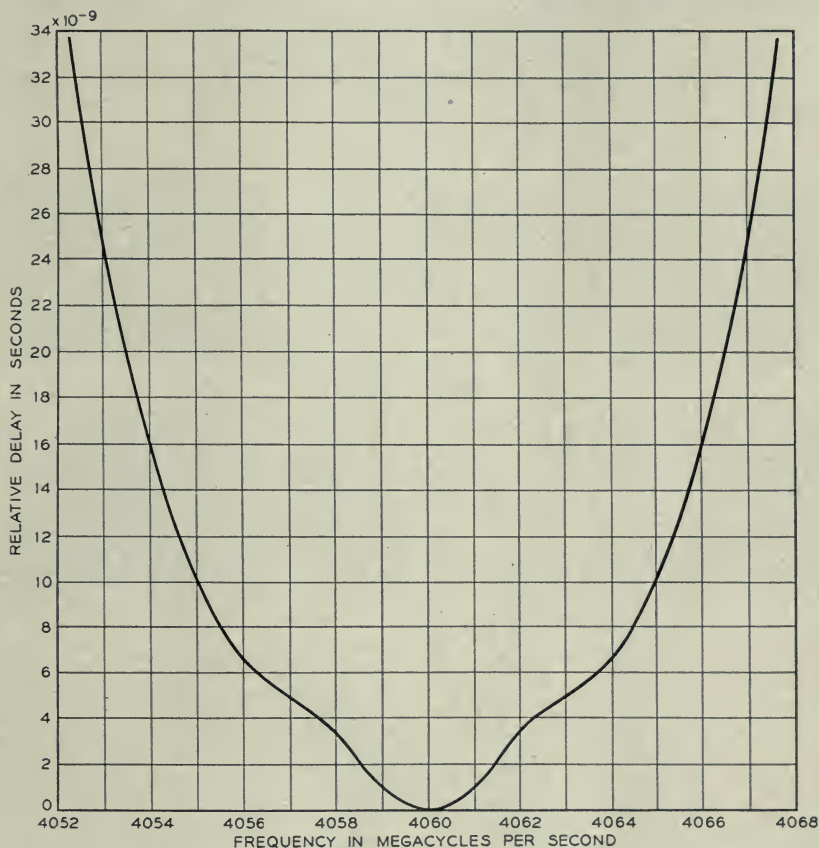


Fig. V-18.—Delay characteristic of RF amplifier.

10, and 30 trips through the IF amplifier. Part of the distortion appearing after one trip through the amplifier was in the viewing oscilloscope. It was difficult to detect distortion from a single trip through the amplifier, but Fig. V-20 demonstrates how it increases with successive trips. The second row of pulse pictures shows the improvement achieved by adding phase equalizers to the circuit. It was particularly noticeable in this test that without equalizers the details in the shape of the pulse after 10 or 30 trips were very sensitive to the exact value of the intermediate frequency.

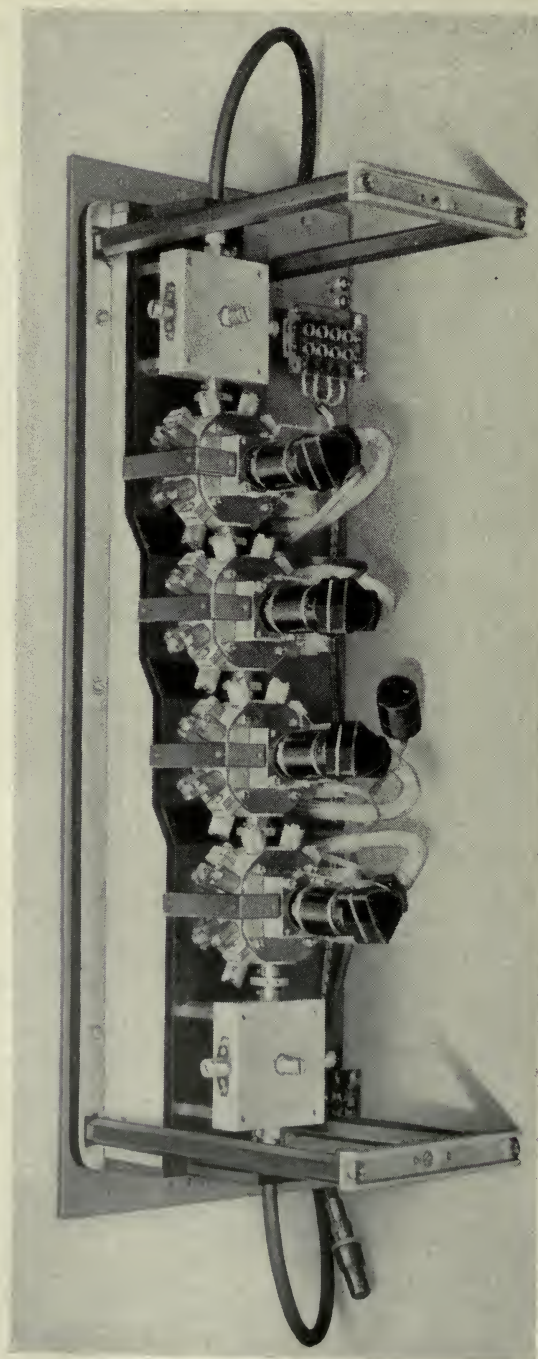


Fig. V-19.—The complete RF amplifier with cover removed.

When the equalizers were added, the pulse shapes shown in the second row were not greatly changed for variations of ± 3 megacycles or more in the intermediate frequency.

In concluding this section it should be noted that when the various components were connected together to form an IF-type repeater amplifier, it was found that the amplitude and phase characteristics added as expected and resulted in a satisfactory amplifier with only a few millimicroseconds variation in delay and only a few tenths of a db variation in amplitude over a 10-megacycle band. Nevertheless, the equipment is very complicated.

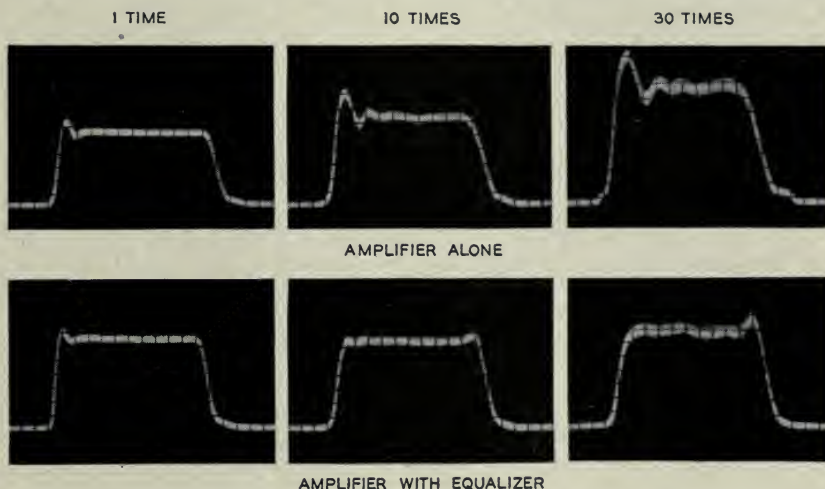


Fig. V-20.—Results of circulated pulse tests on IF amplifier.

A straight-through radio frequency amplifier repeater is still to be desired and, no doubt, further research will eventually produce such an amplifier.

VI. THE COMPLETE REPEATER*

In the preceding section it has been pointed out how the various components that make up the repeater amplifier were added together without the introduction of additional distortion. The antennas, filters and amplifiers which go to make up one of the complete repeaters shown in Fig. II-1 were designed with the same ease of interconnection in mind. However, as will be discussed below, the length of the waveguide lines used for these connections has an important bearing on the distortion introduced.

The large amount of equipment in the repeater amplifier makes it desirable, from the maintenance standpoint, to locate the amplifier near the ground and to provide towers for the antenna where antenna elevation

* Prepared by D. H. Ring.

is necessary. This means relatively long transmission lines between the antenna and the rest of the repeater. If exact termination of the line by the antenna impedance or by the filter input impedance were possible no distortion would be produced, but in general there will be a slight mismatch and a corresponding reflection of energy at each of these junctions which

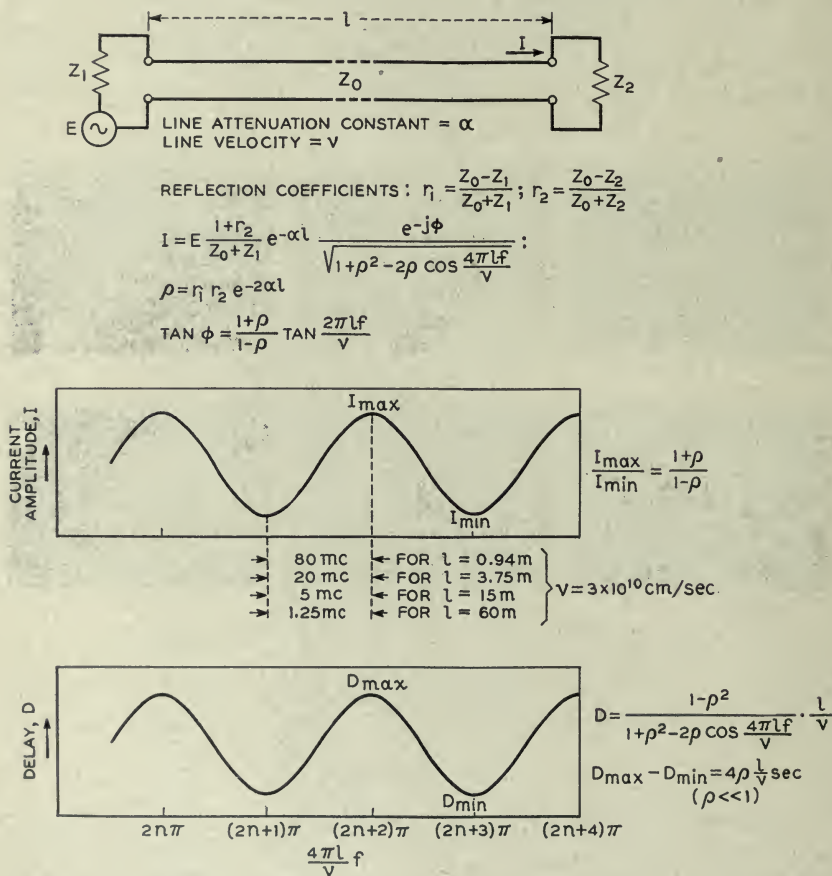


Fig. VI-1.—Effect of long lines on the amplitude and delay distortion of a repeater.

will produce variations in the amplitude and delay of the signal throughout the desired band. These variations may be greater than those produced anywhere else in the repeater.

The type of distortion originating in this way is illustrated by Fig. VI-1. In this figure there is represented an antenna of impedance Z_1 connected by a line of length l and characteristic impedance Z_0 , to a load, Z_2 . In actual practice this load is the impedance presented by the filter to the line from

the antenna. The variations produced in both the amplitude and delay of the signal currents are shown by the curves in the figure where relative change in the characteristics in question are shown as ordinates and values of the quantity $\frac{4\pi\ell f}{v}$ are shown as abscissas.

It will be seen from these curves that the amount of variation over a given band is a function of the line length. With short lines the amount of variation over a band 10 mc wide can be kept very small, but with lines about 150 feet in length there may be three or four full cycles of variations in a band 10 megacycles wide. Table A gives some typical values of the variations. Values are given for those degrees of mismatch which would produce standing wave ratios of 1, 1.4 and 2 db at the junctions. It will be seen that for lines between 100 and 200 feet in length the variations in some cases are greater than the limits achieved in the multistage amplifiers and other com-

TABLE A

$SWR_1 = SWR_2$ in db ($\alpha = 0$).....	1	1.4	2
$r_1 = r_2$ ($\alpha = 0$).....	0.058	0.082	0.115
Max. Amplitude Variation in db.....	0.058	0.116	0.233
Max. Delay Variation in 10^{-9} seconds.....			
$\ell = 100$ feet.....	1.33	2.67	5.33
$\ell = 200$ feet.....	2.67	5.33	10.66

ponents of the system. Furthermore, it is usually impractical to compensate or tune out these variations because they are functions of the electrical length of the transmission lines and subject to change with temperature, frequency, and other small mechanical and electrical changes. It would appear that with present techniques this may be one of the most serious sources of distortion in a long relay system.

VII. CONCLUSION

Various phases of our work on microwave repeater circuits have been discussed. Such a circuit, made up of several repeaters as described in the last section, may be looked upon as a four-terminal network having specified amplitude and delay characteristics and should be suitable for the transmission of any signals for which the characteristics are adequate whether they be television signals or those of one of the various forms of multiplex telephony. It is outside the scope of this paper to discuss the uses to which such a repeater circuit might be put or the terminal equipment that any particular service might require.

The New York-Boston repeater circuit has been built to provide the experimental field trials necessary to answer many remaining questions which

deal with the performance of an actual circuit. The results of these trials will be reported in a separate paper.

ACKNOWLEDGEMENTS

The work which has been discussed in this report of necessity involved the combined efforts of many individuals. In addition to those already mentioned, practically every other member of the staffs of the Deal and Holmdel Radio Laboratories has made valuable contributions to the work, as have also many in other departments of the Bell Telephone Laboratories.

The development of the components of the New York-Boston circuit has been under the direction of Mr. Gordon N. Thayer.

In particular, we wish to acknowledge the support and stimulating advice of Dr. R. Bown under whose general direction the work progressed.

The Measurement of Delay Distortion in Microwave Repeaters*

By D. H. RING

Measuring equipment is described which is capable of measuring delay distortion of the order of 10^{-9} seconds in a wide band microwave television relay repeater. Two measuring circuits are discussed. The first is a circuit for measuring the relative phase shift versus frequency from which the delay distortion may be computed. The second circuit gives the delay directly from a single measurement. The measuring equipment is designed to work in the intermediate frequency range from 50 to 80 megacycles, but by applying suitable conversion equipment measurements can be made at microwave frequencies.

THE successful transmission of broadband television and pulse signals over any communication circuit depends upon the preservation of the complex wave shapes of the original transmitted signals. Fourier analysis tells us that a complex signal wave can be resolved into a spectrum of frequencies with certain amplitude and phase relationships. It is well known that the amplitude relationships of all essential frequencies in this spectrum must be substantially preserved. It is equally important that the phase relationships of the essential frequencies should be preserved. The instantaneous value of the received signal is the vector sum of the instantaneous amplitudes of all the component frequencies. Therefore, if the relative phase of some frequency component is changed by 180° the sign of its contribution to the output is reversed, and it is clear that a closer approximation to the original signal could be obtained by suppressing this frequency component rather than permitting it to contribute negatively to the output.

It can be shown¹ that the relative phase relations of the component frequencies in a complex signal wave will be preserved if the phase shift in passing through a circuit is a linear function of the angular frequency. That is

$$\beta = T_0\omega + n\pi \quad (1)$$

where T_0 is a constant and n an integer. Distortion of the transmitted signal will occur if T_0 is not constant over the essential frequency band of the signal. We shall not be interested in distortion due to n not being an integer² since this case does not occur in carrier circuits where the phase shift at carrier frequency, rather than the phase shift at zero frequency, is

* Presented at National Convention of I. R. E., New York City, March 4, 1947.

¹ Phase Distortion and Phase Distortion Correction, S. P. Mead, *B.S.T.J.*, April 1928, 199-201.

² Phase Distortion in Telephone Apparatus, C. E. Lane, *B.S.T.J.*, July 1930, 494-496.

the reference point for phase. Departure of β from the linear relationship given by (1) is the phase distortion in the circuit.

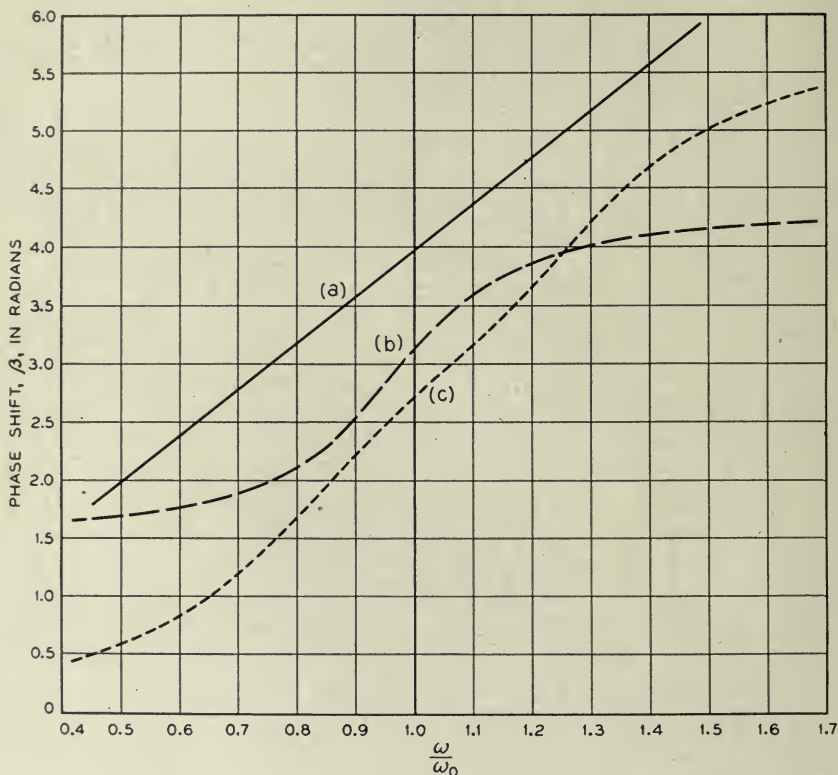


Fig. 1—Typical phase shift curves for various types of circuits
 a. A Transmission line terminated in its characteristic impedance.
 b. A single tuned circuit.
 c. Two tuned circuits with approximately critical coupling.

The time of transmission³ or the delay in passing through the circuit is obtained by differentiating (1):

$$\text{Delay} = \frac{d\beta}{d\omega} = T_0 \text{ seconds} \quad (2)$$

Variation in T_0 with frequency is the delay distortion in the circuit.

Figure 1 shows some typical phase curves, and Fig. 2 shows the corresponding delay curves. In each figure curve (a) represents an ideal distortionless circuit with linear phase and constant delay such as a simple transmission line. Curves (b) are obtained for single resonant circuits,

³ T_0 has also been called the group delay and the envelope delay.

and curves (c) for coupled double tuned circuits. It is felt that the delay curves of Fig. 2 are easier to interpret and give a better physical picture of the distortion resulting from phase variations than the phase curves of Fig. 1. Therefore, most of the following discussion will be in terms of delay rather than phase.

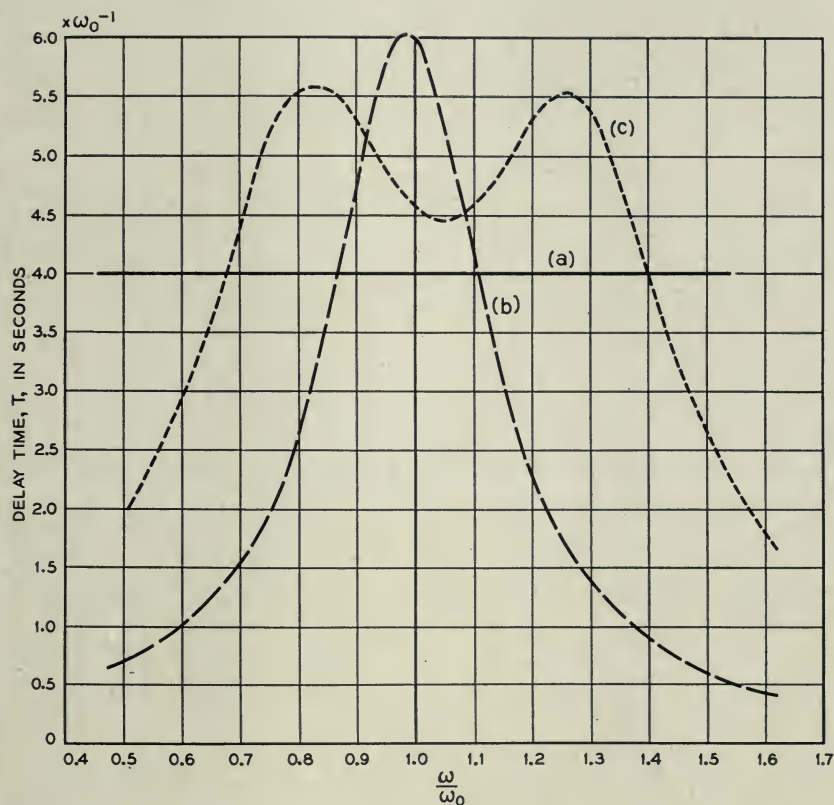


Fig. 2—Typical delay curves for various types of circuits

- a. A transmission line terminated in its characteristic impedance.
- b. A single tuned circuit.
- c. Two tuned circuits with approximately critical coupling.

Thus far we have considered the general case and stated that the delay must be constant over the essential frequency band for distortionless transmission. It should be noted, however, that when delay distortion is present different types of signals may be affected differently. For instance, in the case of an amplitude modulated carrier of angular frequency ω_0 , Fig. 2, we note that if the delay curves have arithmetic symmetry about ω_0 , then the sidebands at $\omega_0 \pm \Delta\omega$ will suffer the same delay and therefore will add

in phase upon demodulation. If the delay distortion is not symmetrical about ω_0 the sidebands will not add in phase upon demodulation, and the demodulated output will suffer both amplitude distortion and some delay distortion which differs from the delay distortion at both $\omega_0 + \Delta\omega$ and $\omega_0 - \Delta\omega$. In the case of frequency modulation dissymmetry introduces harmonics in the demodulated output. A detailed discussion of these effects is beyond the scope of this paper, but we may note that in general a true picture of the delay distortion in carrier circuits is not readily obtained by observing the demodulated output and that an unsymmetrical delay distortion is particularly undesirable in carrier circuits.

PRINCIPLES OF DELAY MEASUREMENT

Delay cannot be measured directly on a steady state basis with a single test signal in the simple manner in which amplitude response is measured. Instead, the phase shift through the unknown network must be measured at two adjacent frequencies and the delay, or slope of the phase shift, computed from the relation

$$T = \frac{\Delta\beta}{\Delta\omega} \quad (3)$$

Figure 3 illustrates the computation of the average delay in the interval $\Delta\omega$ from two phase measurements.

The steady state phase shift of an unknown network can be measured by using the basic circuit shown schematically in Fig. 4.⁴ This is essentially a bridge circuit in which the phase shift in the unknown is balanced by an equivalent calibrated phase shifter. The phase comparator is some kind of device which will give an indication of a known relationship between the phases of the signals arriving over the unknown path and the known reference path.

The exact form of suitable components and circuit arrangements for applying the basic method of Fig. 4 to a particular delay measuring problem is largely determined by the order of magnitude of the delay to be measured and accuracy desired. In the case of microwave television repeaters we are interested in video bands of the order of 5 mc wide. As a rough estimate we might say that the highest frequency in the band should not be shifted more than one quarter period from its normal phase position. In the case of linear delay distortion or a parabolic phase-frequency characteristic, one quarter period for 5 mc is 0.05 microseconds. In a repeater system with 50 repeaters this yields a tolerable systematic delay distortion of 10^{-9} seconds per repeater. Therefore we conclude that in developing

⁴ Measurement of Phase Distortion, H. Nyquist and S. Brand, *B.S.T.J.*, July 1930, 526-527.

repeaters for this service an accuracy of better than 10^{-9} microseconds in measuring the relative delay over a band of frequencies will be desirable. The phase shift, $\Delta\beta$ in Fig. 3, which corresponds to a delay of 10^{-9} seconds is a function of the measuring interval $\Delta\omega = 2\pi\Delta f$. If Δf is small $\Delta\beta$ will

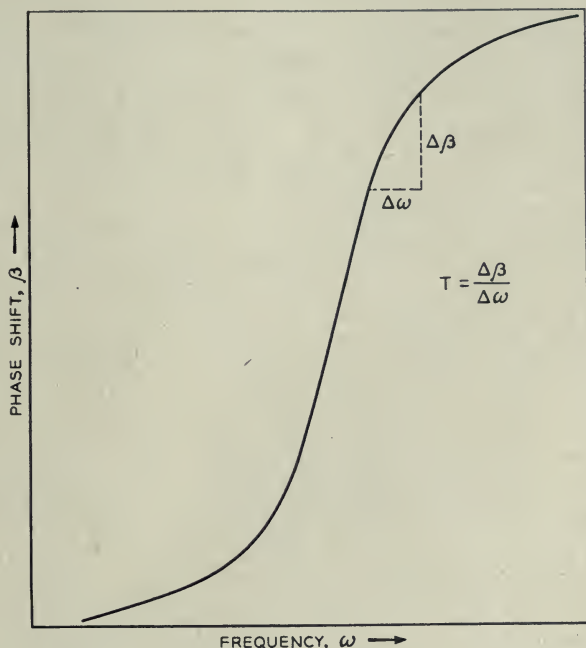


Fig. 3—Factors involved in calculating the delay of an electrical circuit

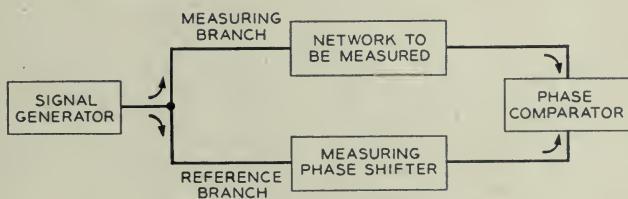


Fig. 4—Basic circuit for measuring the phase shift in a network

be small and difficult to measure. If Δf is large the average slope measured will not be the true slope in the center of the interval. For a 5-megacycle video signal the intermediate and radio frequency bands of interest will be in excess of 10 megacycles wide. These considerations led to a choice of 1 megacycle as a reasonable value for Δf . If $\Delta f = 1$ megacycle a phase shift $\Delta\beta = 0.36^\circ$ will result from a delay of 10^{-9} seconds.

Two circuits for measuring delay distortion will be described. The first is a phase measuring circuit which, in principle, is an adaptation of Fig. 4 to practical operation at intermediate frequencies. The second circuit is a modification in which two frequencies differing by Δf are fed through the circuit under test simultaneously in such a way that the difference in the phase shift at the two frequencies is measured. This arrangement permits the calculation of the delay from a single measurement and is, therefore, a delay measuring circuit.

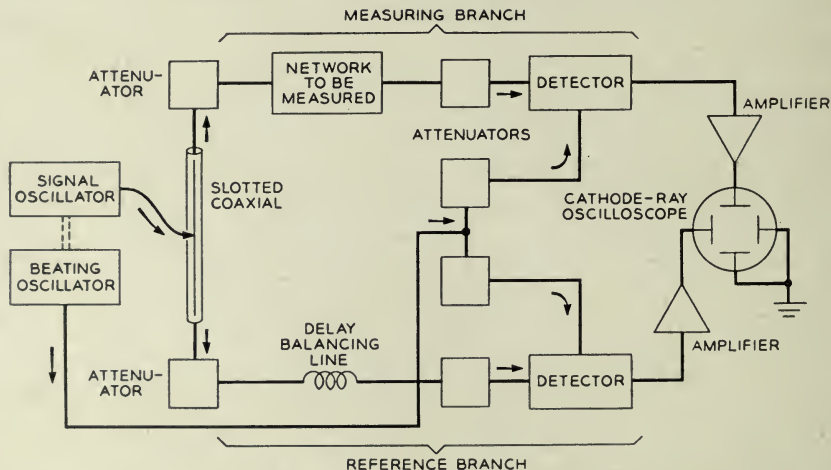


Fig. 5—Schematic circuit for the precise measurement of phase shift in the intermediate frequency range.

PHASE MEASURING CIRCUIT

Figure 5 shows a schematic diagram of a phase measuring circuit which has been found to be suitable for precision measurements of wide-band circuits in the intermediate frequency range of 50 to 80 megacycles. It is a double detection system with an intermediate frequency of one megacycle. The test signal oscillator and the beating oscillator are ganged to a single control and adjusted to track so that they maintain a difference of approximately one megacycle throughout their tuning range. The test signal is fed to a sliding contact on a section of air dielectric coaxial transmission line. The signal divides at this point to feed the test branch and reference branch.

The sliding tap on the coaxial line provides a high precision measuring phase shifter if the coaxial line is well terminated at each end. The relative change $\Delta\beta$ in phase of the signals at the two ends of the line when the

slider is moved a distance Δx is two times the phase shift corresponding to the movement of the slider at the working frequency or

$$\Delta\beta = \frac{720f\Delta x}{c} \text{ degrees}$$

where c is the velocity of light.

For $\Delta x = 0.1$ centimeter and $f = 65$ megacycles, $\Delta\beta$ is 0.156 degrees.

The measuring branch and reference branch feed through the network to be measured and a balancing line respectively to identical detectors and one megacycle amplifiers. The amplifiers are connected to the plates of a cathode ray oscilloscope which is used as a phase comparator. A cathode

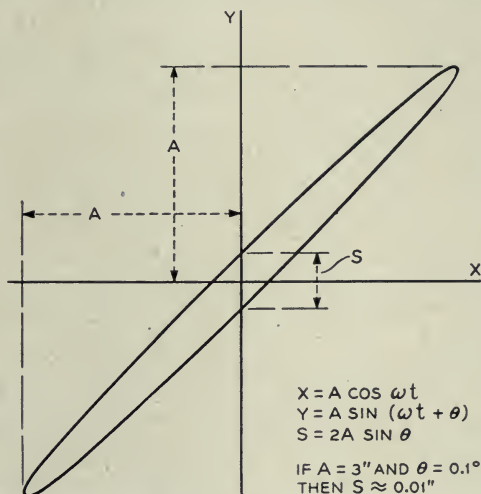


Fig. 6—Calculation of the sensitivity of a cathode ray oscilloscope as a phase comparator.

ray oscilloscope has the advantage that the phase comparison with this instrument is independent of either the relative or absolute amplitudes of the two signals. A straight line on the oscilloscope always indicates that the two voltages are in phase (or phase opposition) regardless of the relative amplitudes, which determine the slope of the line. The sensitivity, of course, is a function of the amplitudes. Figure 6 illustrates how the sensitivity of an oscilloscope phase comparator can be calculated. If each signal alone produces a 6-inch deflection, then 0.1 degree phase difference produces an opening of the pattern of 0.01 inches, which is sufficient to be detected on a "rocking" or in-out basis.

It is obvious that a circuit of this type measures the difference in the phase shifts of the two branches. The measured phase shift is, therefore, the absolute phase shift in the measuring branch less the phase shift in the ref-

erence branch. The balancing line shown in the reference branch of Fig. 5 can be adjusted in length so that it balances out the average delay in the measuring branch. The measured remainder will then be the distortion in the measuring branch, since a good transmission line does not have phase or delay distortion. The use of a balancing line in the reference branch simplifies measurements by reducing the range of movement of the slider, and it greatly decreases the errors in the calculation of the delay distortion due to inaccuracies in the measuring interval Δf .

The simplified diagram of Fig. 7 will be used to show how the delay of the unknown network can be calculated. With the signal oscillator set at frequency f_1 the slider is adjusted until the signals at C and E are in phase

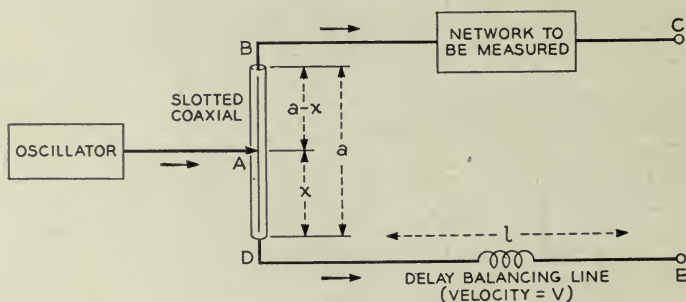


Fig. 7—Simplified circuit electrically equivalent to the circuit of Fig. 5.

as indicated by a straight line on the oscilloscope, and the corresponding distance x_1 is measured. The relationships between the phases of the signals at the various points in the circuit are:

$$\begin{aligned}\phi_C &= \phi_A - \phi_{AB} - \phi_{BC} \\ &= \phi_A - \frac{2\pi f_1}{c} (a - x) - \beta\end{aligned}\quad (4)$$

$$\begin{aligned}\phi_E &= \phi_A - \phi_{AD} - \phi_{DE} \\ &= \phi_A - \frac{2\pi f_1 x_1}{c} - \frac{2\pi f_1 l}{v}\end{aligned}\quad (5)$$

$$\phi_C = \phi_E$$

where β is the phase shift in the unknown network at frequency f_1 . Solving for β ,

$$\beta = \frac{2\pi f_1}{c} \left[\frac{lc}{v} + 2x_1 - a \right]\quad (6)$$

The signal oscillator frequency is then changed to $f_2 = f_1 + \Delta f$ and the slider readjusted to the position x_2 that again makes $\phi_c = \phi_E$. Then as above

$$\beta' = \frac{2\pi f_2}{c} \left[\frac{\ell_c}{v} + 2x_2 - a \right] \quad (7)$$

where β' is the phase shift in the unknown network at frequency f_2 . The delay in the network is, from (3)

$$T = \frac{\beta' - \beta}{2\pi\Delta f} \quad (8)$$

Substituting (6) and (7) in (8) yields

$$T = \frac{1}{c} \left[\left(\frac{\ell_c}{v} - a \right) + 2x_2 + \frac{2f_1}{\Delta f} (x_2 - x_1) \right] \quad (9)$$

The first term in T is a constant of the set-up and can be dropped when the delay distortion only is desired. The second term is small and can often be neglected. The third term gives the major part of the difference between the delay of the reference path and the delay of the network.

It has been found that the slider position can be easily reset to ± 0.05 centimeters for each frequency. This would mean a maximum error of ± 0.1 centimeter for the difference of two readings which corresponds to an error in T of about $\pm 0.4 \times 10^{-9}$ seconds. However, this reset accuracy will not be realized in overall accuracy unless a number of precautions are observed in setting up the circuit of Fig. 5. In order to avoid stray coupling the two branches of the system must be carefully shielded from each other. At least 60 db net attenuation must be provided between the detectors via the path through the balancing line, phase shifter and unknown network, and 40 db attenuation in the path via the BO supply lines. All attenuators, plugs, etc., must have voltage standing wave ratios of less than about 1.015 and the detectors and amplifiers in each branch must have identical phase shifts over the range of variation of the IF due to imperfect tracking of the oscillators.

DELAY MEASURING CIRCUIT

The circuit of Fig. 5 is basically a phase measuring circuit. It can be rearranged as shown in Fig. 8 to yield a circuit that will read delay directly from a single setting of the slider. In Fig. 8 the signals from the two oscillators are both sent through the circuit to be measured, and both are sent through the reference branch. Any delay in either path will alter the relative phases of the two signals in that path and this change in relative phase shift will be passed on to the beat note formed in the detectors. If

the delays in the two paths are different a relative phase shift proportional to the difference will appear in the two beat notes. This phase change can be measured with a phase shifter in the beat note circuit or with a phase shifter which varies the relative phase of the two signals fed to one branch

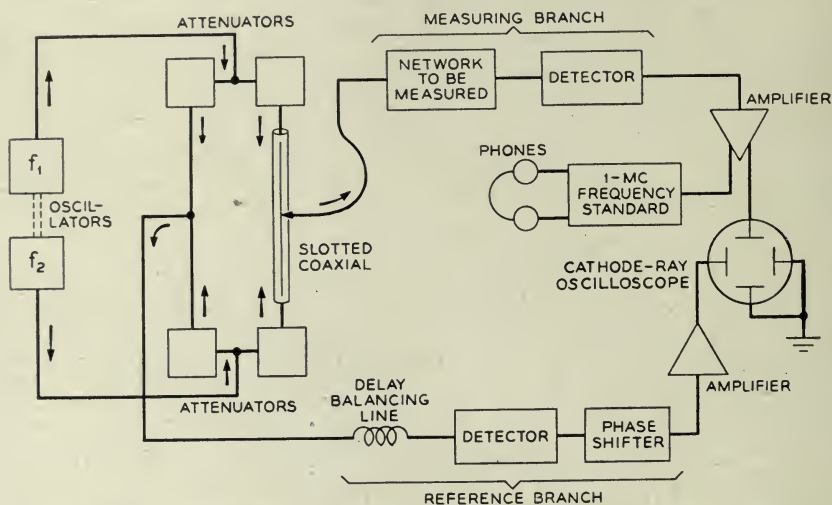


Fig. 8—Schematic circuit for the precise measurement of delay in the intermediate frequency range.

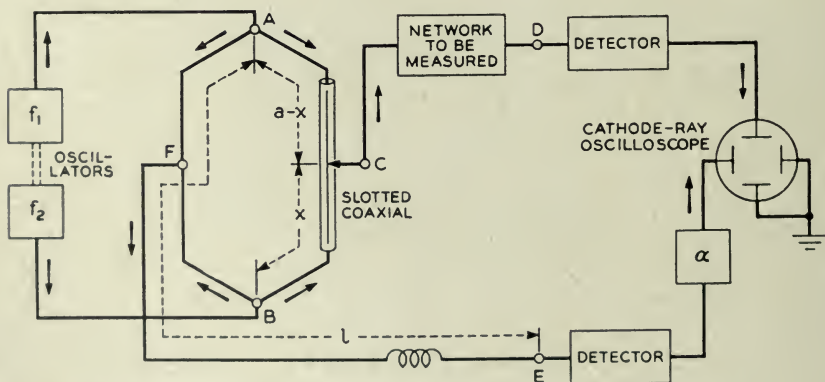


Fig. 9—Simplified circuit electrically equivalent to the circuit of Fig. 8

as compared with the relative phase of the two signals fed to the other branch. Figure 8 shows how the coaxial slider can be used for the latter type of measurement.

The simplified diagram of Fig. 9 will be used to show how the delay of the unknown network can be calculated in terms of this circuit. It will be assumed that the electrical lengths of the paths AF and BF are equal so

that this length can be regarded as part of the balancing line ℓ . Then at point E we have

$$\phi_E = \phi_A - \phi_{AE} = \phi_A - \frac{2\pi f_1 \ell c}{v}$$

$$\phi'_E = \phi'_B - \phi'_{BE} = \phi'_B - \frac{2\pi f_2 \ell c}{v}$$

The primes indicate phase shifts at frequency f_2 . The phase ϕ''_E of the beat note, $\Delta f = f_2 - f_1$, at E is

$$\phi''_E = \phi'_E - \phi_E = \phi'_B - \phi_A - \frac{2\pi \ell c}{v} \Delta f \quad (10)$$

Similarly, at point D

$$\phi''_D = \phi'_B - \phi_A - \frac{2\pi f_2 x}{c} + 2\pi f_1(a - x) - (\beta' - \beta) \quad (11)$$

If x is adjusted so that the Δf 's in the two branches are in phase at the oscilloscope then

$$\phi''_D = \phi''_E - \alpha \quad (12)$$

where α is the difference in the phase shifts in the two beat note circuits. Substituting (10) and (11) in (12) and solving for $(\beta' - \beta)$ yields

$$(\beta' - \beta) = \frac{2\pi}{c} \left[\frac{\ell c \Delta f}{v} - x \Delta f - f_1(2x - a) + \frac{\alpha c}{2\pi} \right] \quad (13)$$

The delay in the network is

$$T = \frac{(\beta' - \beta)}{2\pi \Delta f}$$

$$= \frac{1}{c} \left[\frac{\ell c}{v} + \frac{\alpha c}{2\pi \Delta f} - f_1 \frac{(2x - a)}{\Delta f} - x \right] \quad (14)$$

The first two terms in (14) are independent of x and f_1 and therefore are of interest only for absolute measurements. Relative measurements may be made without evaluating these constants. The last two terms are functions of x and yield the change in delay as f is varied. It is usually most convenient to adjust α and ℓ so that the average delay in the measuring branch is given by (14) with

$$(2x - a) = 0 \quad (15)$$

In general this will minimize the variation of the slider and make the delay distortion in the measuring branch roughly proportional to the slider

movement. This is helpful in judging the effect of adjustments of the unknown network. The condition (15) corresponds to the optimum condition for the circuit of Fig. 5 which requires that the average delays of the two branches be equal. In the circuit of Fig. 8, the condition (15) can be realized by varying either ℓ or α . If α is made zero and (15) is fulfilled by adjusting ℓ , errors due to variations in Δf are minimized. However, if Δf is held sufficiently constant, the balancing line ℓ can be omitted entirely and a 360° variable phase shifter introduced in the beat note circuit of one branch to vary α .

The measuring interval Δf is determined by the difference in the two signal oscillators. Since Δf has an important influence on the measurement, oscillator tracking cannot be relied upon to maintain Δf with sufficient accuracy. A one-megacycle crystal frequency checker has therefore been included in the equipment as shown in Fig. 8. A sample of the signal from one of the amplifiers is compared with the crystal oscillator and a trimmer on one of the oscillators adjusted for zero beat before measuring each point. When T is so large that a balancing line is not practicable, as in the case of loop circuits including radio paths or long transmission lines, Δf must be held constant to about 1 part in 10^6 . This has been accomplished in a modification of this equipment built by Messrs. W. J. Alberheim and J. P. Shafer of these laboratories, by deriving the two measuring frequencies from a crystal oscillator.

In order to obtain the absolute value of T , the constants ℓ , v , and α must be known. The value of ℓ may be found from the physical length of the line; α can be measured by measuring the movement of the slider required to rebalance the circuit when the connections to the two detectors are reversed. The absolute delay is particularly sensitive with respect to the difference between $2x$ and a . If a is measured as accurately as possible, then the exact setting of the slider corresponding to condition (15) can be found by reversing the connections of the slider at points A and B in Fig. 9. In this way a reference value of x may be found which will yield accurate results in spite of a small error in the value of a .

Successful operation of this circuit depends on low standing waves throughout and upon sufficient padding for satisfactory isolation of the oscillators. It will be noted that in the analysis it was assumed that f_1 reached the slider via the path AC, Fig. 9. There must be an attenuation for f_1 in the path AFBC sufficient to render the signal traversing this path negligible. Similar unwanted paths exist for f_2 from B to C and also for f_1 and f_2 to point F. Attenuation inserted as shown in Fig. 8 can be arranged to make these unwanted signals negligible.

RADIO FREQUENCY MEASUREMENTS

The frequency range of a particular measuring equipment using the circuits of Fig. 5 or Fig. 8 is determined only by the range of the oscillators and the range over which the plugs and jacks and attenuators operate with sufficiently low reflection coefficients. While this range is greater than the IF range encountered in microwave repeaters, it does not include the actual microwave frequencies. However, it has been found that microwave components can be measured satisfactorily by using the circuit shown in Fig. 10. In this circuit the measuring equipment is operated at

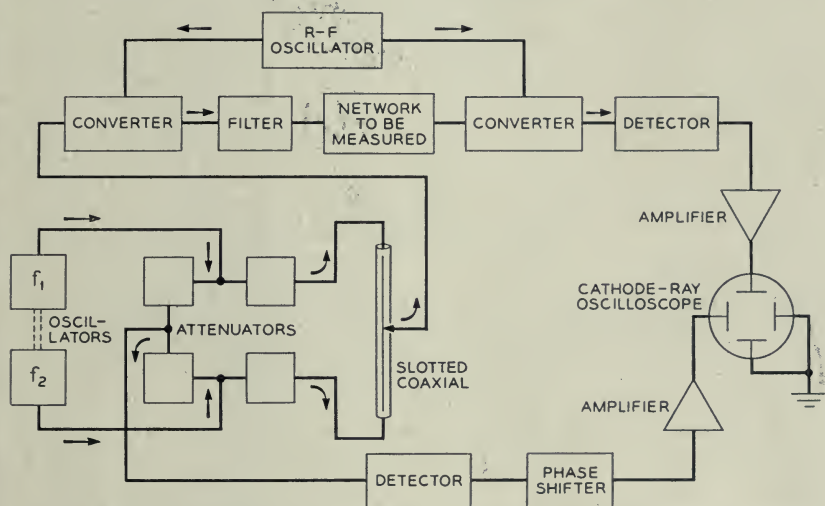


Fig. 10—Method of using the intermediate frequency measuring circuit of Fig. 8 for the measurement of radio frequency networks.

IF, and the reference branch is unaltered. The measuring branch signal is fed first to a converter where it is beat up to the desired microwave frequency. The converter is followed by a filter which eliminates the beating oscillator frequency and one of the beat frequencies. The filter output is then applied to the microwave component under test. The output of this is fed to another converter and converted back to IF by beating with the same oscillator that was used in the first converter. This process removes any variations in phase due to variations in the beating oscillator if the connections from the beating oscillator to each converter are of equal electrical length.

Since considerable extra equipment is included in the measuring branch in Fig. 10, it will usually be necessary to make a calibration run with the

device to be measured removed in order to eliminate any possible delay distortion in the converters and filter. Figure 10 uses the measuring

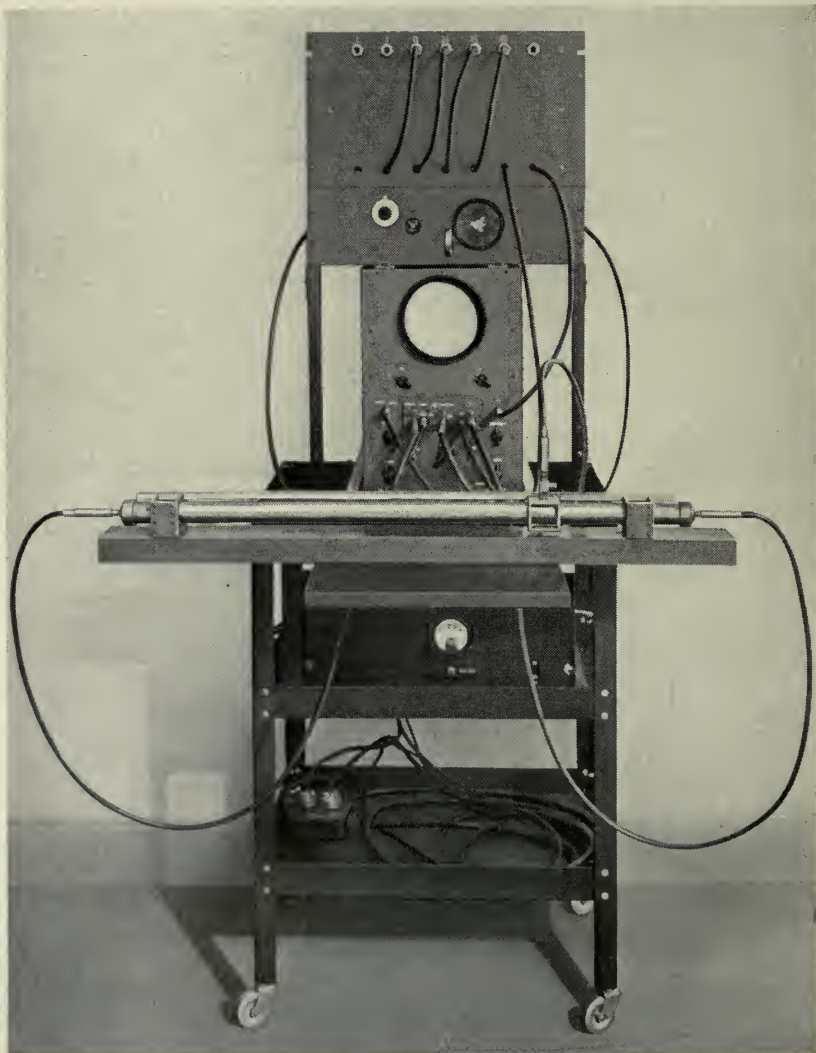


Fig. 11—Photograph of apparatus used in the measuring circuits shown in Figs. 5 and 8.

circuit of Fig. 8 rather than that of Fig. 5 because it was found that small variations in the transit time in microwave amplifiers cause small variations in the phase shift which are the same at all frequencies in the band. These

variations cause changes in the relative phase of the two successive measurements required when using the circuit of Fig. 5 which do not represent changes in the delay. In effect the circuit of Fig. 8 makes the two phase measurements simultaneously, and thus eliminates effects due to variations of phase with time.

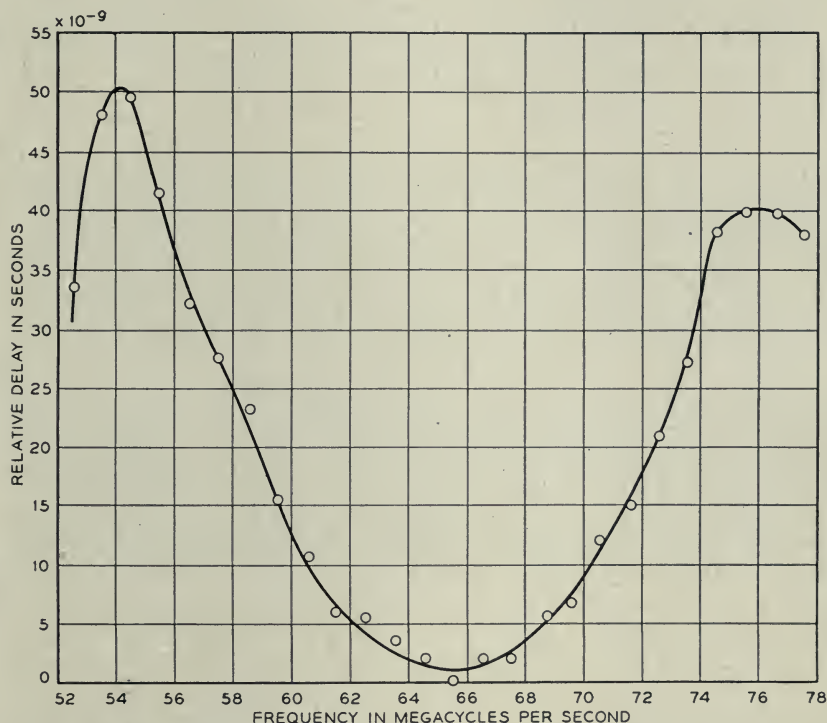


Fig. 12—Measured curve of the relative delay of an intermediate frequency amplifier

RESULTS

Figure 11 shows a photograph of the delay measuring equipment which has been described. With the aid of patch cords and switches that have been included in the equipment this apparatus can be set up according to either Fig. 5 or Fig. 8. The ganged oscillators are on the panel above the oscilloscope box. The box contains the dividing attenuators, detectors, amplifiers and oscilloscope. A separate power supply is required for the oscillators and the output stages of the amplifiers. A number of different lengths of flexible coaxial cable are mounted on the panel above the oscillators and arranged so that various lengths of balancing line can be conveniently obtained by patching. The coaxial phase changer can be seen on the shelf in front of the oscilloscope.

Figure 12 shows the measured relative delay of a typical intermediate frequency amplifier. This amplifier has a substantially flat amplitude response over a band of about 12 mc centered on 65 mc. The delay distortion over the 10 megacycle band from 60 to 70 mc is about 10×10^{-9}

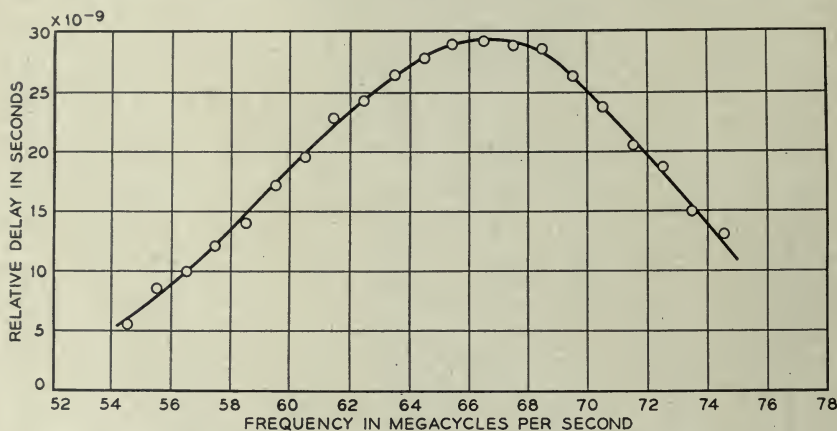


Fig. 13—Measured relative delay of an experimental delay equalizer

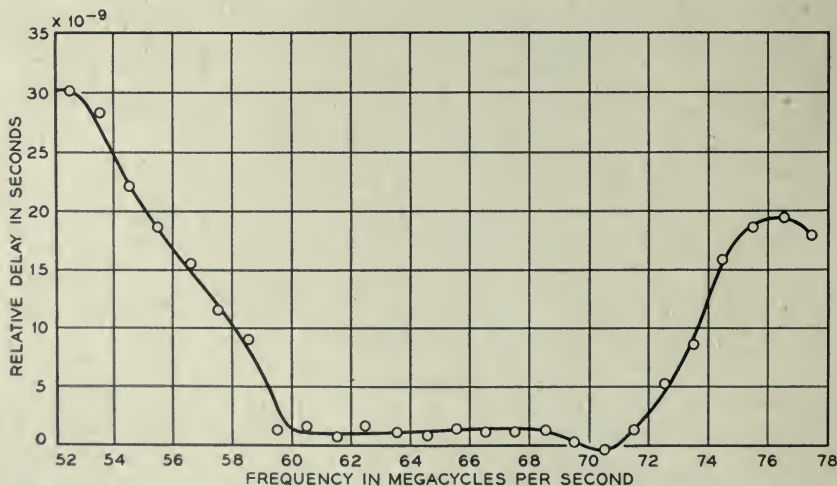


Fig. 14—Measured relative delay of the amplifier of Fig. 12 plus an equalizer

seconds. Figure 13 shows the measured relative delay of a bridged T phase equalizer which can be used to compensate for the distortion in the amplifier. Figure 14 shows the measured relative delay for the amplifier of Fig. 12 and an equalizer measured together. The equalizer has reduced the delay distortion over the 10 mc band from about 10^{-8} to 10^{-9} seconds. These

measurements were made with an early model of the measuring equipment. Smoother curves are obtained with the apparatus shown in Fig. 11.

In Fig. 15 the top row shows the distortion of a square top pulse by the amplifier of Fig. 12 for 1, 10, and 30 trips through the amplifier without the equalizer. The lower row shows a similar set of pictures of the pulse when the distortion was reduced by a phase equalizer as shown in Fig. 14. The improvement due to the elimination of phase distortion is clearly illustrated. These pictures were obtained by the circulated pulse⁵ technique which

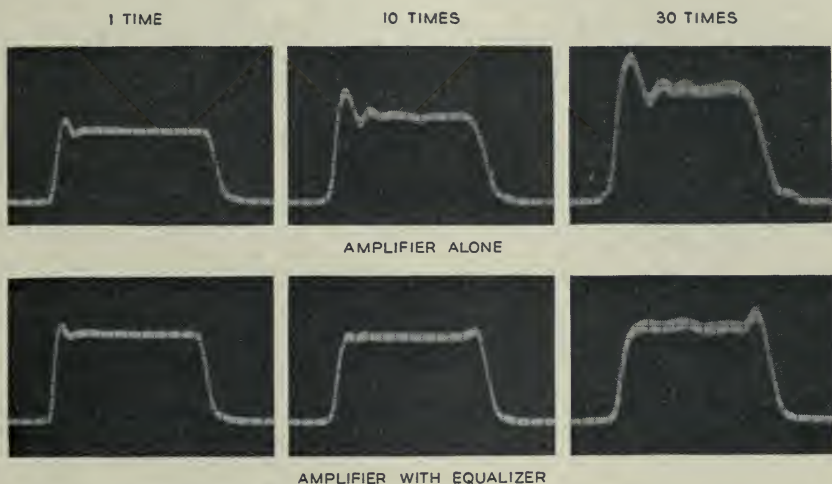


Fig. 15—Oscillograms showing the improvement in the square wave response of the amplifier of Fig. 12 obtained by delay equalization. The numbers above the traces indicate the number of times the test pulse has passed through the amplifier and equalizer.

permits the observation of a pulse after it has passed a number of times through the same amplifier.

CONCLUSIONS

Two measuring circuits have been described which are suitable for measuring the small variations in relative transmission time which are present in wide band microwave repeaters. If sufficient care is exercised in setting up these circuits an accuracy of better than $\pm 10^{-9}$ seconds can be realized in relative delay measurements. The circuit of Fig. 5 measures the relative phase shift as a function of frequency. It has the advantage of requiring less signal power and fewer important parameters for absolute

⁵ Testing Repeaters with Recirculated Pulses, A. C. Beck and D. H. Ring. Proc. I.R.E. Nov. 1947, 1,226-1,230.

delay measurements. The circuit of Fig. 8 measures delay directly with a single measurement and has the advantage of ignoring uniform phase variations with time. It is useful for making relative measurements on circuits with long constant delay times. A significant improvement in the square wave response of carrier amplifiers has been obtained by applying delay equalizers based on measurements made with this equipment.

Frequency Shift Telegraphy—Radio and Wire Applications*

By J. R. DAVEY and A. L. MATTE

Frequency shift telegraphy is described and compared with amplitude modulation telegraphy under various conditions found in radio and wire transmission. Experimental data are given to demonstrate the influence of various design factors on the over-all performance under these conditions. It is shown that the most outstanding characteristic of the frequency shift method is its ability to accept large and rapid changes in signal amplitude. Frequency shift telegraphy thus proves to be of great advantage for use in the H.F. radio range. Frequency shift telegraphy also shows an advantage over amplitude modulation telegraphy with respect to noise. For applications where the level variations are small or slow the advantage of the frequency shift method over amplitude modulation is relatively small.

INTRODUCTION

DURING World War II, single-channel and multichannel frequency-shift radio telegraph systems proved of the utmost importance in providing the Allied Powers with a world-wide automatic printing telegraph network for handling with precision, secrecy and dispatch the unprecedented volume of traffic engendered by a war of global extent. It is expected that the next few years will witness a greatly expanded application of this method of operation by commercial telegraph companies and others interested in long distance telegraphy.

Frequency Shift carrier telegraphy (FS) may be applied to any carrier telegraph circuit, but, as will appear below, it provides particularly striking advantages in H.F. radio transmission. For some other radio frequency ranges and for wire line operation the conditions are such as to limit the advantages of the FS method. The main advantages of the FS over the AM method are a greater ability to accept rapid level changes, which results in better stability and lower distortion, and an improvement in signal-to-noise ratio, which permits a reduction in carrier amplitude. It is therefore of particular importance where automatic printing is desired over H.F. radio circuits. When it is necessary to transmit through very high noise levels, low speed AM signaling with aural reception of an audio beat note remains the superior method.

FS is a form of frequency modulation in which signaling is accomplished by shifting a constant amplitude carrier between two frequencies representing respectively the marking and spacing conditions of the telegraph code. Frequency variations in FS telegraphy correspond to amplitude variations in

*Published in *A.I.E.E. Transactions*, Vol. 66, p. 479, 1947.

AM telegraphy (CW); thus the signal transitions in FS are represented by frequency-time transients, while in the AM case they are amplitude-time transients. Since AM telegraph is the more common system, a discussion of the FS method involves numerous comparisons between the two systems. The merits of a telegraph system must be judged on its ability to combat the various adverse conditions encountered in the transmission medium and in the terminal apparatus. In general these adverse conditions involve variations in amplitude, frequency, and phase of the signals and the presence of extraneous signals and noise.

In the course of the development of a number of FS radio teletypewriter systems, a large amount of information concerning the characteristics and design parameters of such equipment has been obtained. It is the purpose of this paper to abstract therefrom selected data which will furnish a step-by-step comparison of the FS and AM methods. Typical terminal arrangements are described and the effects of varying certain design factors are illustrated by experimental data. Although the material presented applies largely to H.F. radio telegraph, much of it is of a general nature and with proper interpretation applies to other frequency ranges and transmission mediums and to cases in which the telegraph modulated carrier may be a sub-carrier or one of several sub-carriers.

GENERAL DISCUSSION

Sideband Energy Distribution

The difference between FS and AM signals as regards distribution of sideband amplitudes is illustrated by the following two equations for a carrier of frequency $\omega/2\pi$ modulated with unbiased square wave dots of frequency $\rho/2\pi$.

For AM (On-off) keyed carrier of unity amplitude¹:—

$$\begin{aligned}
 e = & 0.5 \cos \omega t + \frac{1}{\pi} [\cos (\omega + \rho)t + \cos (\omega - \rho)t] \\
 & - \frac{1}{3\pi} [\cos (\omega + 3\rho)t + \cos (\omega - 3\rho)t] \\
 & + \frac{1}{5\pi} [\cos (\omega + 5\rho)t + \cos (\omega - 5\rho)t] \dots \text{etc.} \quad (1)
 \end{aligned}$$

For FS keyed carrier of unity amplitude²:—

$$e = \frac{2m}{\pi} \left[\frac{1}{m^2} \sin \left(\frac{m\pi}{2} \right) \cos \omega t \right]$$

$$\begin{aligned} & + \frac{1}{m^2 - 1^2} \cos\left(\frac{m\pi}{2}\right) (\cos(\omega - \rho)t - \cos(\omega + \rho)t) \\ & - \frac{1}{m^2 - 2^2} \sin\left(\frac{m\pi}{2}\right) (\cos(\omega - 2\rho)t + \cos(\omega + 2\rho)t) \\ & - \frac{1}{m^2 - 3^2} \cos\left(\frac{m\pi}{2}\right) (\cos(\omega - 3\rho)t - \cos(\omega + 3\rho)t) \\ & + \dots \dots \dots \end{aligned} \quad (2)$$

phase swings, the modulator would have to be able to produce a steady phase rate of change. A phase modulator capable of performing in this way while not producing undesirable phase discontinuities at the signal transitions becomes rather impractical. For this reason FS telegraphy usually utilizes the direct frequency modulation method. This may conveniently be accom-

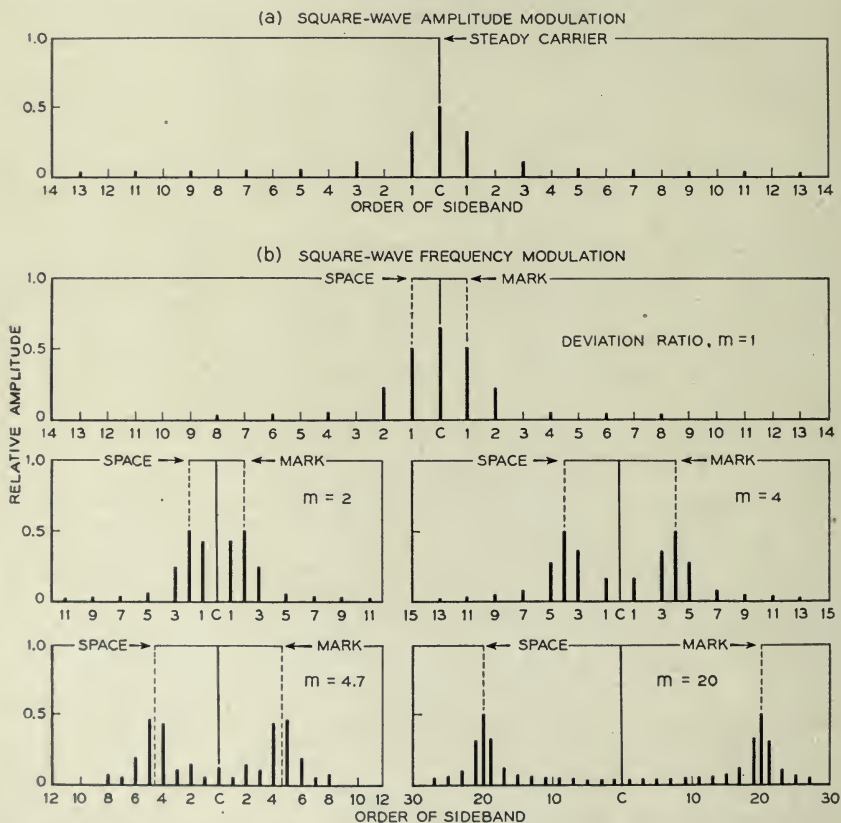


Fig. 1.—Amplitude of sideband components for (a) square-wave amplitude modulation (b) square-wave frequency modulation.

plished by the use of a reactance modulator which, by injecting a reactive component of current into the tuned circuit of the oscillator, varies the resonant frequency thereof. Such a modulator may be made linear so that a frequency shift proportional to the input voltage to the reactance modulator is obtained.

To apply FS telegraph signals to a radio transmitter the regular exciter oscillator is either replaced or modified by an arrangement providing a source of R.F. excitation that can be shifted in frequency in accordance

with the telegraph signal. All the stages are operated with full R.F. excitation continuously to produce a constant amplitude carrier.

As a matter of expediency frequency shift keying has sometimes been provided by switching between two independent sources of carrier current separated in frequency by the desired shift. In such a case the frequency transitions involve sudden phase discontinuities of random values. This results in the instantaneous frequency swinging considerably outside the steady-state mark and space frequencies. If the band is wide, such as is the case in a radio transmitter, there results a very broad sideband radiation capable of causing severe interference to adjacent channels. If the band is narrow, as might be the case where sending filters are employed, the interference is eliminated but the amplitude transients resulting from the sudden phase shifts are capable of producing considerable distortion.

Restriction of Transmitted Band

As seen above, square-wave modulation results in a wide spread of sideband components which are of sufficient amplitude to interfere seriously with adjacent channels unless greatly attenuated. The transmitted band may be restricted either by the use of a band-pass filter centered about the carrier frequency or by a low-pass filter to suitably shape the modulating wave form. Band-pass filters are usually used if the power level is low and the frequency low enough to permit suitable filter construction. For multi-channel systems the use of band-pass filters also permits efficient paralleling of the transmitting channels. For radio transmitters with a transmitted power measured in kilowatts, and with a frequency of several megacycles which is frequently changed to suit best the prevailing conditions, shaping of the modulating wave is the more practical method of restricting the transmitted band.

Insufficient attention has been given in the past to the envelope shape of the signals from on-off keyed radio transmitters. With the ever increasing crowding of frequency assignments it becomes more and more important to restrict the emission of unnecessary sidebands arising from keying. The envelope shape in on-off keying may be controlled by properly shaping the modulating grid or plate voltage wave. It is important that the stages following the keyed stage or stages be nearly linear, otherwise the wave shaping will be largely destroyed. In the case of frequency shift keying, on the other hand, the wave shaping is preserved after passage through class C amplifier or multiplier stages, and these may be operated for maximum efficiency. The greater ease of producing and maintaining the desired wave shaping, so necessary for close frequency spacing of channels, is one of the outstanding advantages of frequency shift keying.

APPARATUS

Typical FS Exciter for Radio Telegraph

A typical FS exciter arrangement such as is often used with radio telegraph transmitters is shown in Fig. 2. A d-c. telegraph wave, after suitable shaping, is caused to frequency-modulate an intermediate frequency of 200 kc. which, in turn, amplitude-modulates a radio frequency from a crystal-controlled oscillator. The upper sideband of this latter modulation is an FS signal and is selected and amplified sufficiently to drive the first amplifier or multiplier stage of the transmitter. The 200-kc. oscillator is frequency-modulated by a reactance modulator which, by feeding a leading or lagging

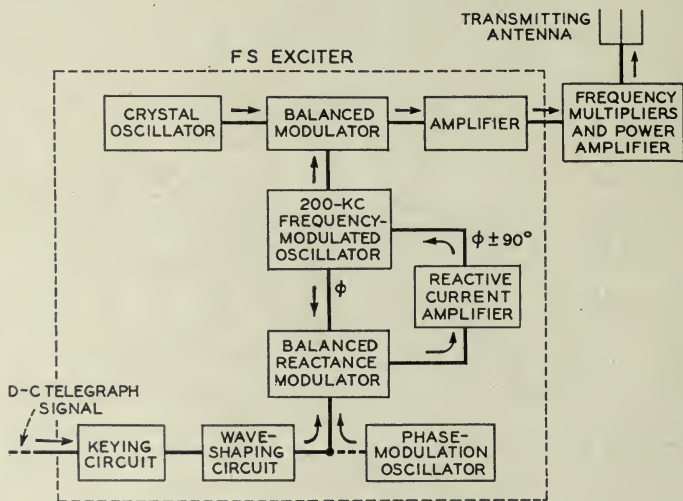


Fig. 2.—Block diagram of a typical FS transmitter.

quadrature component of current into the oscillator tuned circuit, decreases or increases the frequency. By operating the reactance modulator within its linear range the frequency shift wave form is made the same as the d-c. telegraph wave form into the modulator. A d-c. amplifier stage, designated "keying circuit", is provided to furnish a modulating wave effectively isolated from amplitude and wave front variations of the incoming telegraph signals. The d-c. telegraph signals may be polar or neutral and are often obtained from a tone demodulator unit which allows keying from a remote point by V.F. telegraph. The amount of frequency shift is adjusted by an amplitude control in the quadrature feed-back path to the 200 kc. oscillator. The shift may thus be varied continuously, or in definite steps to allow for subsequent frequency multiplications, by suitable attenuation controls. Controlling the shift in this manner keeps the instabilities of the reactance

modulator a constant percentage of the frequency shift, which would not be the case if the shift were adjusted by varying the amplitude of the modulating wave. The use of a balanced instead of an unbalanced reactance modulator minimizes variation of the mean frequency and also allows the shift to be varied without affecting the mean frequency.

The frequency-shift signal transitions are wave-shaped, to restrict sideband radiation, by means of a low-pass filter in the d-c. telegraph signal path to the reactance modulator. The low-pass filtering is made adjustable to accommodate a range of signaling speeds. Frequency-versus-time wave

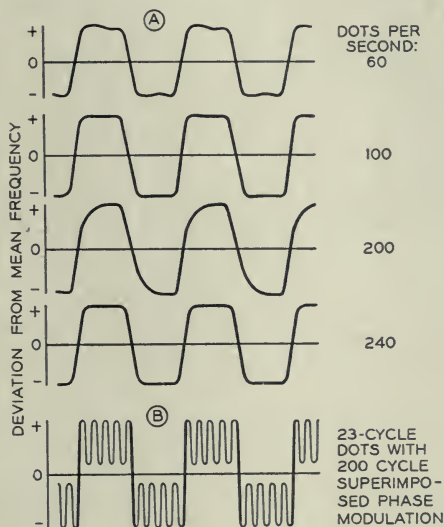


Fig. 3.—Frequency shift keyer output wave forms. (a) Low-pass filtering adjusted to produce similar wave shapes at dotting speeds of 60, 100, 200, and 240 cycles. (b) 200 cycle phase modulation superimposed on a 23 dots per second signal.

forms from an exciter of the type described are shown for several keying speeds in Fig. 3a. The effect of wave shaping on the sideband components in the R.F. output of such an exciter is shown in Fig. 4.

Phase modulation may readily be added to the signal in this type of exciter by superimposing the desired sine wave modulating frequency on the telegraph signal wave input to the reactance modulator as indicated in Fig. 2. Figure 3b shows the keyer output wave form with superimposed phase modulation. The use of this type of phase modulation is considered later.

To obtain optimum results in FS radio telegraph transmission and to allow close spacing of channels, a high degree of frequency stability is necessary. An over-all frequency stability of ± 100 cycles is desirable in a system using a value of frequency shift between 500 and 1000 cycles. A frequency

shift exciter of the type described above, with the crystal oscillator and 200 kc. FS oscillator located in a temperature-controlled oven, usually has a frequency stability such that the mean R.F. carrier frequency may be held to within ± 50 cycles up to frequencies of 20 mc over ordinary periods of operation on any one frequency. One of the advantages of this type of exciter is that small inaccuracies in crystal frequencies may be compensated for by adjusting the mean frequency of the 200 kc. oscillator.

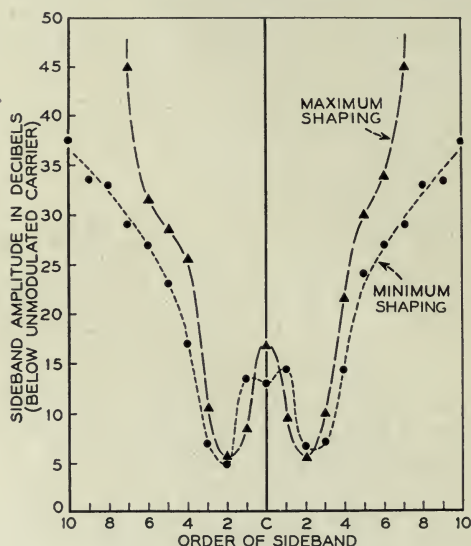


Fig. 4.—Effect of shaping FS transitions on amplitude of radiated sidebands. 100 dots per second and 500 cycle frequency shift. The maximum and minimum wave shaping conditions correspond to those used with 60 and 240 dots per second respectively in Fig. 3(a).

Receiving Terminal Arrangements

Typical receiving terminal arrangements are shown in the block diagrams of Fig. 5. Up to point "a" in the arrangements the FS and AM systems are identical, being of the usual H.F. superheterodyne type. The output from the second frequency-conversion stage may be either in the audio range or at a considerably higher frequency such as 50 kc. Following the second converter is a band-pass filter (shown at "b" in Fig. 5) which determines the final over-all band width before demodulation. The two systems differ only in the method of demodulation. The AM (on-off) signals are amplified and rectified to give a d-c. telegraph signal. The FS signals are amplitude limited and passed through a frequency discriminating network and then rectified to give a d-c. telegraph signal. Beyond this point the two systems are again identical. The d-c. signals pass through a low-pass filter to remove

carrier ripple and higher frequency noise components and are then amplified to a suitable level to operate automatic recording or printing apparatus. The d-c. signals may also be used to modulate an audio frequency so as to pass the signals to a remote point by multichannel voice-frequency carrier telegraph methods.

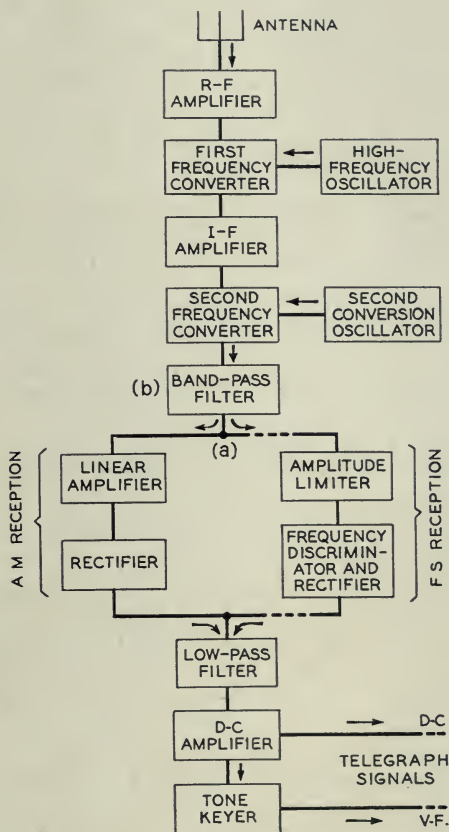


Fig. 5.—Block diagram of a typical receiving arrangement for either AM or FS telegraph signals.

The radio receiver portion of the terminals up to point "a" should be designed to have low noise and good selectivity. Extreme H.F. oscillator stability is necessary for either system if narrow band width operation is to be maintained without constant attention. An over-all frequency stability of ± 50 cycles is desirable for the receiving terminal over a period of 6 to 8 hours. Sufficient selectivity and amplifier capacity should be provided at all points to prevent overloading by unwanted signals or loss of automatic

gain control. In the following discussion those portions of the terminals beyond the second frequency converter will be given major attention.

Experimental Transmitting and Receiving Arrangements

For the laboratory transmission studies described in the following sections the transmitter and receiver were located nearby and connected together by means of an amplitude modulator and associated with various sources of noise designed to simulate quantitatively and under controlled conditions the variations which would be encountered in the actual medium.

Throughout the tests 7.42 unit start-stop signals were used unless otherwise stated, and the speed was 60 words per minute (23 dots per second). Their peak distortion and bias were measured on a cathode-ray tube telegraph distortion measuring set.

An exciter of the type shown in Fig. 2 was used as a source of signals. A frequency of 6.4 mc. was employed, with the radio receiver connected to the exciter output through an amplitude modulator. This modulator was an electronic circuit permitting amplitude modulation of a frequency-shift signal to produce unequal mark and space amplitudes. This modulator was also used to amplitude-modulate a single frequency for the AM portions of the measurements.

A temperature-limited diode together with a two-stage tuned amplifier was used as a source of thermal noise centered around 6.4 mc. A polar relay driven by 60-cycle a-c and arranged to produce sharp polar impulses from the discharge of small capacitances connected to its contacts was used as a source of impulse noise. The noise level was adjusted by an attenuator and mixed with the 6.4 mc. carrier of the exciter. A minimum amount of wave shaping was used, so that the modulation may be considered as having been essentially square-wave.

Receiving Arrangements

The experimental data submitted in the following discussion was obtained from reception through a laboratory setup essentially like that shown in Fig. 5. The radio receiver proper was a commercial type of H.F. superheterodyne. The output of the second frequency converter was in the audio-frequency range, which enabled the use of various band-pass filters at "b" of the type used in voice-frequency telegraph systems. The amplitude limiter was effective over an input range of -60 dbm* to above $+20$ dbm. The pass-band characteristics of the radio receiver and of the several band-pass filters used in position "b" are shown in Figs. 6 and 7 respectively. Unless otherwise stated, the frequency shift signals were centered about

* The symbol dbm signifies "db referred to one milliwatt".

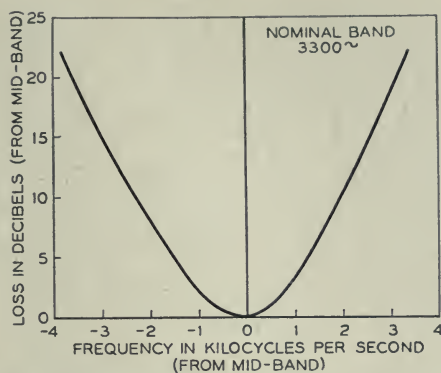


Fig. 6.—I.F. selectivity characteristic of radio receiver.

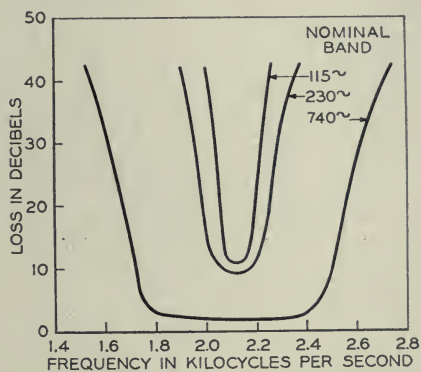


Fig. 7.—Attenuation versus frequency characteristics of bandpass filters shown at (b) in figure 5.

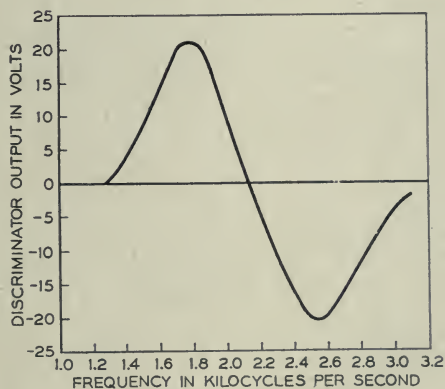


Fig. 8.—Linear discriminator characteristic.

2125 cycles and were demodulated by a linear discriminator centered about 2125 cycles as shown in Fig. 8. The characteristics of the low-pass filtering are shown in Fig. 9. These low-pass filters were adjusted by oscillographic observation of the signal wave form and had cut-off characteristics giving very little characteristic distortion³. The d-c. amplifier was a high-gain nonlinear type designed so as to have a square-wave output having transitions established by the passage of the demodulated voltage wave through a narrow amplitude range. The amplitude and wave front slope of the demodulated wave thus had no effect on the output wave form and could not affect the distortion measuring equipment.

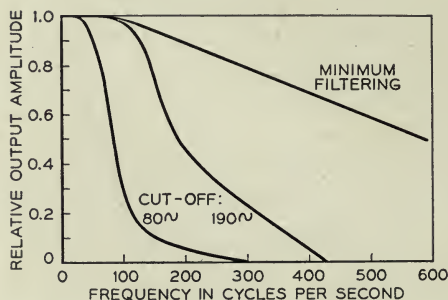


Fig. 9.—Attenuation versus frequency characteristics of low-pass filters.

EXPERIMENTAL RESULTS

Band Width Before Demodulation

The band width before demodulation determines the amount of noise and interference which is to be accepted along with the desired signal and thus largely determines the signal-to-noise condition at the antenna at which the system fails to receive intelligence. The band width at this point (point "a" in Fig. 5) also limits the signaling speed capabilities of the system. In the following experimental data the values of band width were measured between the points of 6 db loss above that at midband.

Effect on Signaling Speed

For both methods of signaling considered here the band width must be at least twice the maximum signaling speed in dot-cycles per second but it is found that signal distortion rises rapidly for a band width less than three times the maximum signaling speed, and that a factor of at least four times is indicated for a system which is to have reasonably low distortion with any margin of safety. The signaling speed capability of a given band width is nearly the same for FS signals as for AM (on-off) signals. In Fig. 10 is

shown the over-all signaling frequency response to FS and AM signals for a nominal band width at point "a" of about 740 cycles. It will be noted that the FS method is but slightly inferior and that both systems fail at a frequency of approximately one-half the band width.

Effect on Noise:

The effect of *thermal noise* on distortion for bandwidths of 115, 230, and 740 cycles is shown in Figs. 11 and 12. The rms. noise-to-carrier ratios

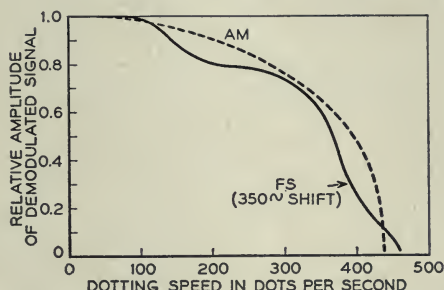


Fig. 10.—Overall frequency response of a 740-cycle band to AM and FS signals.

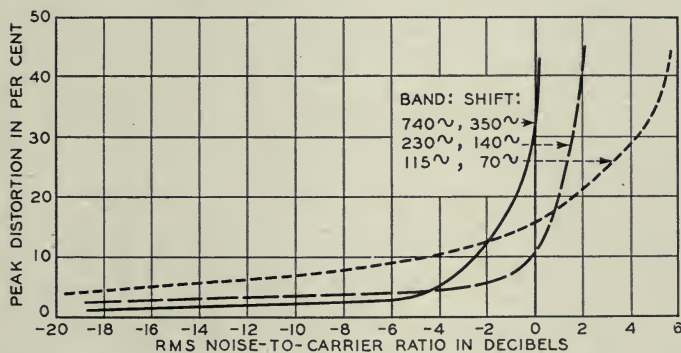


Fig. 11.—Peak distortion versus thermal noise for FS transmission—80-cycle cutoff low-pass filter.

indicated in the figures were measured in the 3300-cycle band of the radio receiver in all cases. The actual noise-to-carrier ratio existing in the transmission band used was lower and may be obtained from the following table. The values of correction are from rms. measurements.

PASS BAND AT POINT "a"	CORRECTION TO BE MADE
1920 cycles	-2.3 db
740	-6.5
230	-11.6
115	-14.3

For *AM signals*, Fig. 12, varying the bandwidth has little effect when the telegraph signal distortion is less than 15%. Although the wider bands accepted more noise power, this added noise merely produced high-frequency noise components which were removed by the low-pass filter. This added noise does, however, cause the peak noise to exceed the signal amplitude at a lower noise-to-carrier ratio and cause failure before that for a narrower band condition.

In the case of *FS signals*, Fig. 11, changing the bandwidth and the frequency shift simultaneously, and in approximately the same proportion, alters the whole distortion characteristic. At low noise levels a wider band with a greater frequency shift gives an improved signal-to-noise condition. However, as the noise level is increased the wider band causes the peak noise

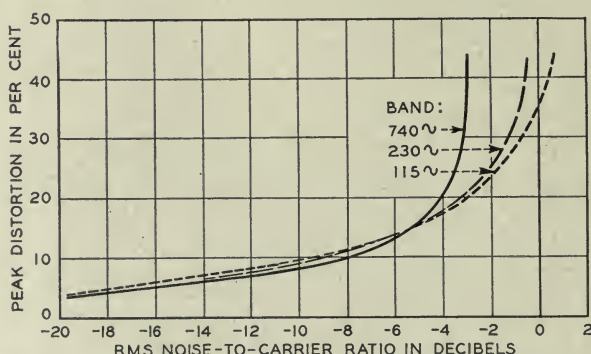


Fig. 12.—Peak distortion versus thermal noise for AM transmission—80-cycle cutoff low-pass filter.

to exceed the carrier at a lower noise level than with a narrower band. Thus a change to a wider band gives less distortion at low noise levels and more distortion at high noise levels. This results in a much more sharply breaking distortion characteristic for the wider band. This behavior is typical of frequency modulation systems in general.

Although the noise actually passed by the 740-cycle band filter was approximately 8 db above that passed by the 115-cycle band filter, the difference in the failure points (35–40% distortion) for the two bandwidths will be seen to be only about half this amount in db for both AM and FS. This phenomenon is typical of carrier telegraph systems when compared at the same signaling speed. This means that it is not particularly beneficial to decrease band width to obtain lower distortion under high noise conditions. The main reason for narrow bands is for more economical use of frequency space.

As to the comparison between *FS* and *AM*, the *FS* method has an advan-

tage of 2.5 to 4.5 db at a distortion of 35% to 40%, corresponding to the selector failure point of the usual teletypewriter. From a signal-to-noise standpoint it is thus seen that the gain in changing to the FS method is approximately equal to the resulting increase in average transmitted power of about 3 db. A comparison at a lower distortion such as 15% shows an advantage of 4 to 6 db. At a still lower distortion the 740-cycle band, because of the higher deviation ratio, shows an improvement of over 10 db. In this region the slopes of the curves make accurate comparisons impossible due to the masking effect of other sources of distortion. These large improvements at low noise levels are similar to those associated with wide-band FM broadcast systems. However, in carrier telegraph transmission the criteria are so different that the difference between a nearly perfect

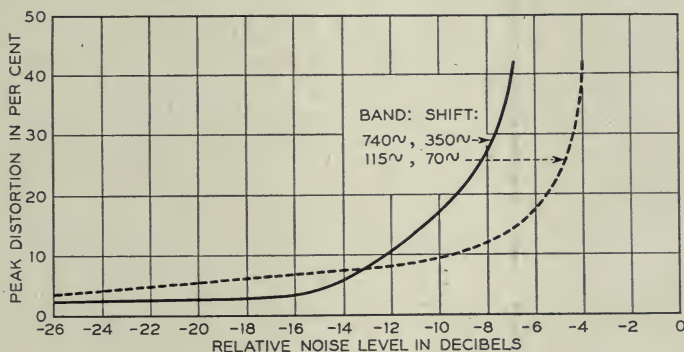


Fig. 13.—Peak distortion versus impulse noise for FS transmission—80-cycle cutoff low-pass filter.

circuit and one of small distortion is not of great importance except when a large number of telegraph sections are to be operated in tandem. From a practical standpoint the improvement in signal-to-noise is not more than about 6 db for equal band widths.

In Figs. 13 and 14 similar characteristics are shown for the case of *impulse noise*. The noise level values of these curves are purely relative since no attempt to measure the peak noise was made. The comparisons between AM and FS, and between different band widths, agree closely with those for thermal noise.

Effect of Limiter on Signal-to-Noise Ratio

The limiter in the FS system is a high-gain nonlinear amplifier which delivers to the frequency discriminating networks an essentially square wave having transitions coinciding with the passage of the instantaneous voltage of the input carrier signal through zero. The limiter thus passes only the

frequency or phase changes of the signal. Noise voltages which are small compared to the signal cause approximately linear phase modulation of the signal and this is passed through the limiter. The amount of frequency deviation thus imparted to the carrier by a given component of noise is proportional not only to its amplitude but also to its frequency separation from the carrier. This gives rise to the so-called "triangular noise spec-

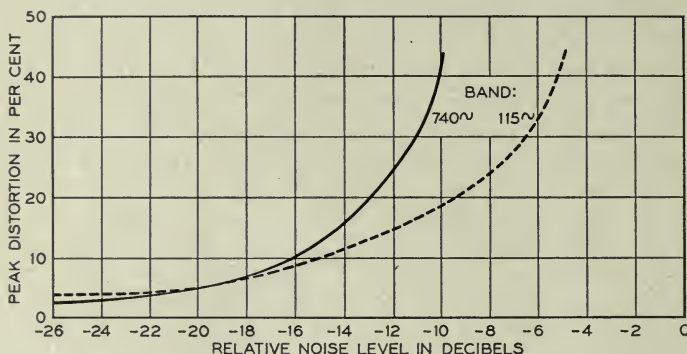


Fig. 14.—Peak distortion versus impulse noise for AM transmission—80-cycle cutoff low-pass filter.

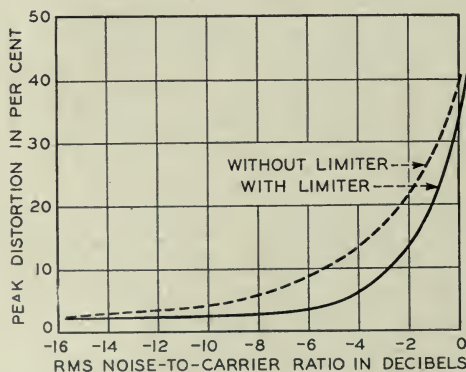


Fig. 15.—Effect of the limiter on distortion versus thermal noise for FS transmission—740-cycle band, 350-cycle frequency shift, 80-cycle cutoff low-pass filter.

trum"⁴ when a linear frequency discriminator is used. If the limiter is removed from the frequency-shift terminal the noise components in phase with the carrier as well as those in quadrature therewith are allowed to reach the frequency discriminating network. For a balanced type discriminator this increases the demodulated noise for small amounts of noise about 3 db. In Fig. 15 is shown the effect of removing the limiter from the circuit. The limiter is seen to have little effect on the failure point, but an improvement of 2 to 4 db is shown in the 5 to 12% distortion region.

Other beneficial effects resulting from the use of a limiter are discussed below under "Level Variations".

Demodulation of Frequency Shift Signals

It is desirable that the frequency discriminating network be of the balanced type having two branches allowing differential combination of the two rectified outputs. This minimizes the response to amplitude modulation not eliminated by the limiter. Two general types of networks have been in common use for FS telegraph. One consists of two bandpass filters centered about the mark and space frequencies respectively and effectively dividing the total band into halves. The other consists of a two-branch network each branch of which has a varying amplitude characteristic extend-

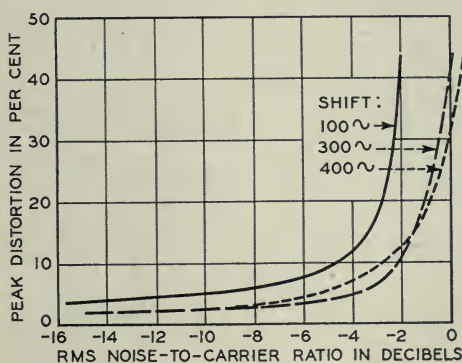


Fig. 16.—Effect of magnitude of frequency shift on distortion versus thermal noise in FS transmission—740-cycle band and 80-cycle cutoff low-pass filter.

ing over the complete transmission band and usually well beyond. The amplitude-versus-frequency characteristics of these two branches have opposite slopes and are of such shape that differential combination of their rectified outputs results in an approximately linear voltage-versus-frequency curve, passing through zero at midband. (Fig. 8)

In Fig. 17 is shown the characteristic of a two-bandpass-filter type of discriminator which was used in early frequency shift terminals. The characteristic is fairly flat near the mark and space frequencies so that this type of discriminator does not produce a triangular noise spectrum. In Fig. 18 is shown the type of discriminator characteristic obtained by the use of two narrow bandpass filters. In this case there is no broad flat region around the mark and space frequencies and an intermediate type of characteristic (approaching the linear type) is obtained.

With the linear type of discriminator the demodulated noise has the well known triangular spectrum and, as illustrated previously, the signaling

speed capability is essentially the same as an AM system of equal band width.

To compare experimentally these two general types of discriminators a 740-cycle band system with linear discriminator and 350-cycle shift was

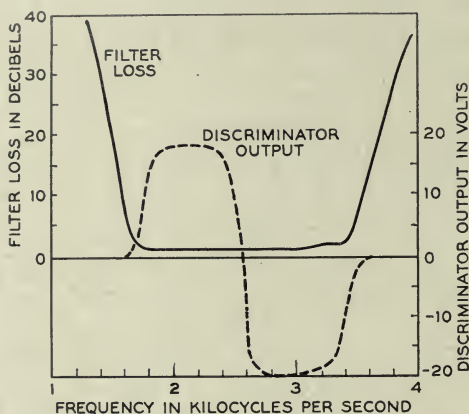


Fig. 17.—Characteristics of 1920-cycle bandpass filter and associated discriminator consisting of two bandpass filters.

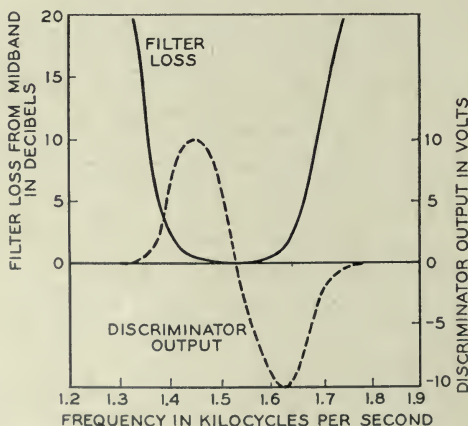


Fig. 18.—Characteristics of 295-cycle bandpass filter and associated discriminator consisting of two bandpass filters.

compared with a system with a bandwidth of about 1900 cycles, a discriminator consisting of two 740-cycle bandpass filters, and an 850-cycle shift. The results are shown in Fig. 19. The two systems are seen to reach failure distortion values at the same signal-to-noise point, with the linear discriminator becoming about 3 db superior at distortions around 5%. A second comparison was made using roughly equal bandwidths. A 295-cycle

band with a discriminator consisting of two bandpass filters and 170-cycle shift was compared with a 230-cycle band with a linear discriminator and a shift of 140 cycles. The results are shown in Fig. 20. The linear discriminator in this case appears to fail slightly sooner but shows a superiority of about 2 db in the 5% distortion region. Due to the rounded character-

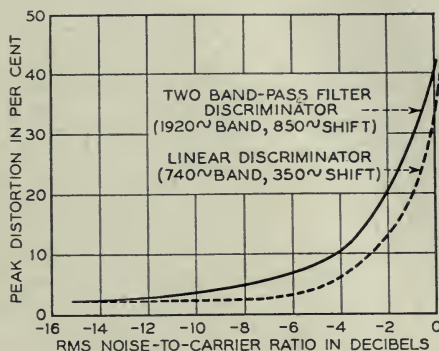


Fig. 19.—Linear discriminator versus two-bandpass-filter discriminator having the same signaling speed capability. Effect of thermal noise on peak distortion—80-cycle cutoff low-pass filter.

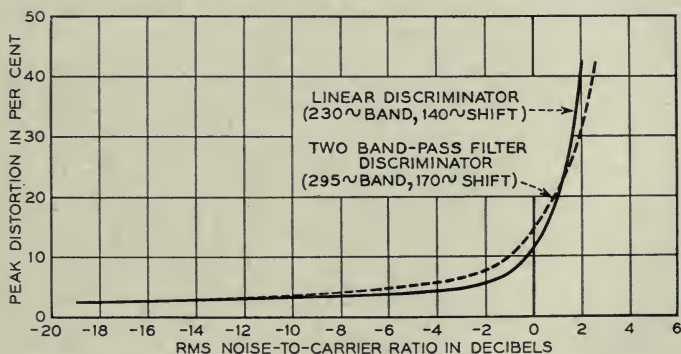


Fig. 20.—Linear discriminator versus two-bandpass-filter discriminator with equal bandwidths. Effect of thermal noise on peak distortion—80-cycle low-pass filter.

istic of the two bandpass filters used as a discriminator in the second comparison (Fig. 18) the difference in discriminators is less than in the previous test. It has been found⁵ that for a given signaling speed capability almost twice the bandwidth is required if a two-bandpass-filter discriminator is used instead of the linear type. This added band width does not appear to cause any loss in signal-to-noise capabilities as to the failure point. The less sharp breaking point, however, makes the linear discriminator superior for moder-

ate and low distortions. For a system occupying a *given bandwidth* the two-bandpass-filter discriminator provides some improvement at the failure point but is still somewhat inferior to the linear discriminator at moderate distortions. More importantly the two-bandpass-filter discriminator impairs the signaling speed capabilities to an extent which depends upon the shape of the cutoff of the filters used.

Bandwidth After Demodulation (Low-Pass Filtering)

In an *Am system* the low-pass filtering after demodulation can, to a large degree, make up for a greater than necessary bandwidth before demodulation. During *marking intervals* the added noise admitted by a wide band causes noise in the demodulated output at frequencies higher than the signaling frequency and this can be filtered out, unless the noise is so great as to over-modulate the carrier. During *spacing intervals* there is no carrier and hence only the noise is rectified. Added noise admitted by a wide band causes not only higher-frequency components in this rectified noise, which may be filtered out, but also an increase in the d-c. component. This tends to cause marking bias of the received signals as the noise level increases.

In an *FS system*, where the carrier is present continuously, the added noise from a wider band produces high-frequency noise components in the demodulated output which can be filtered out by the low-pass filter if the noise level is low. As the noise level increases there are short intervals when the noise envelope exceeds the carrier. The action of the limiter is to give preference to the greater signal, in this case the noise, and since the noise will appear to the discriminator as a carrier fluctuating around mid-band as a center, the demodulated output momentarily dips toward zero. As the noise increases, the duration and frequency of these holes in the signal increase. The low-pass filter, by excluding frequencies considerably in excess of the maximum signaling speed, prevents these holes in the signal from producing false or extra transitions in the telegraph signal output. The low-pass filtering, however, cannot prevent the true transitions from being displaced by this type of noise component since the signal is obliterated momentarily. Its most important function is to prevent a breakup in the signal output until a fairly high distortion is reached. For noise peaks exceeding the carrier the low-pass filter of an AM system also serves much the same purposes.

In Fig. 21 is shown the effect of changing the bandwidth of the low-pass filter in an FS and in an AM system in the presence of thermal noise. The effect of a narrower low-pass filter is seen to consist mainly in shifting the breaking point toward a higher noise level. Similar characteristics for the case of impulse noise are shown in Fig. 22.

Magnitude of Frequency Shift in Relation to Bandwidth

A frequency-shift transient in a band of given width has a wave shape much like that of an amplitude transient in the same band provided the shift is symmetrical and not over 50% of the bandwidth.⁶ If the frequency shift approaches the total width of the band the transient is of such shape as to

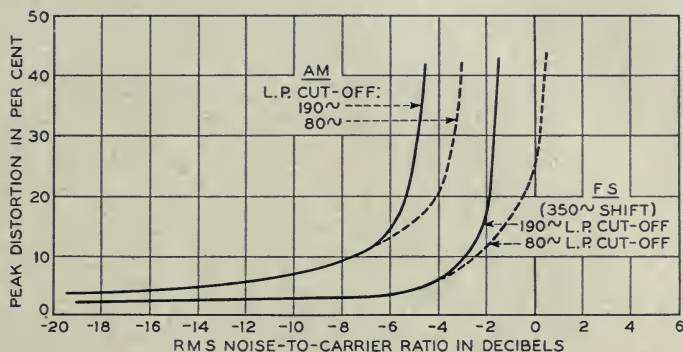


Fig. 21.—Effect of the low-pass filter cutoff frequency on distortion in the presence of thermal noise—740-cycle band.

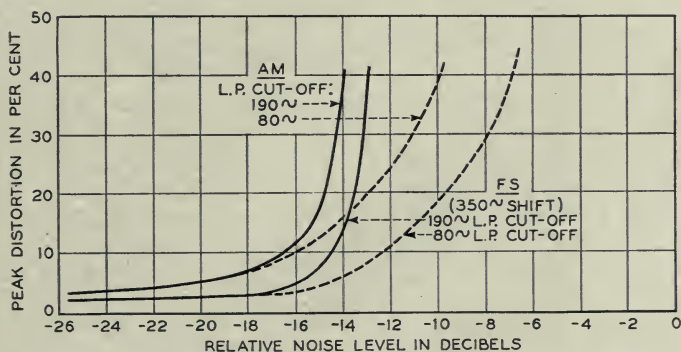


Fig. 22.—Effect of the low-pass filter cutoff frequency on distortion in the presence of impulse noise—740-cycle band.

cause distortion and to make the system more susceptible to noise. A very small shift results in a low amplitude of demodulated signal, which is more readily distorted by noise and biased by frequency drifts. It is of interest, however, that the signal-to-noise ratio of the demodulated signal is not proportional to frequency shift for high noise conditions. As described before, noise peaks tend to reduce momentarily to zero the output from a balanced discriminator. The amplitudes of the dips or holes are thus about one-half the demodulated signal amplitude for any value of shift. This tends to maintain a constant signal-to-noise condition and this characteristic

is illustrated by Fig. 16 in which the breaking point with a 100-cycle shift occurs only 2 db before that with a 400-cycle shift although the difference in actual signal amplitude is 12 db. Figure 23 shows the effect on signal-to-noise ratio of progressively varying the frequency shift while the bandwidth

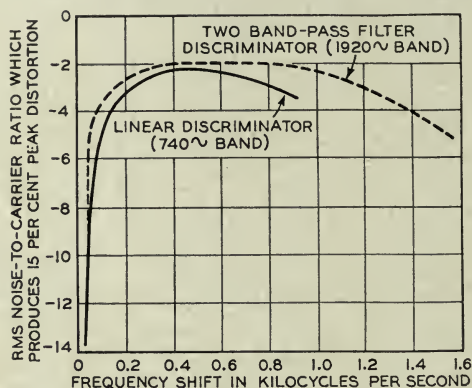


Fig. 23.—Effect of magnitude of frequency shift on distortion produced by thermal noise—80-cycle cutoff low-pass filter.

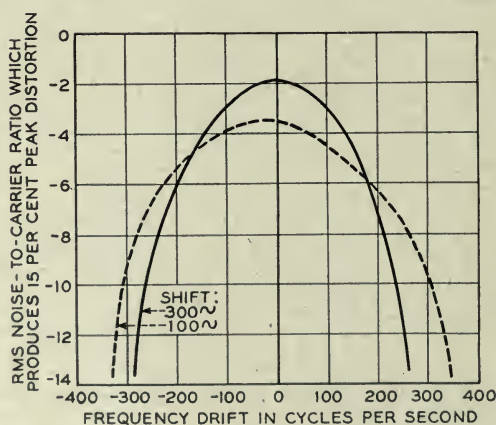


Fig. 24.—Effect of frequency drift on distortion produced by thermal noise in FS transmission—740-cycle band and 80-cycle cutoff low-pass filter.

is kept constant. There is a fairly broad region in which the signal-to-noise ratio is little affected by the amount of frequency shift. For optimum results a frequency shift of 50% to 60% of the band width is indicated, and thus has been used in most of the experimental results given herein.

Frequency Instabilities

Frequency drift of the carrier input to the receiving terminal in general causes biased signals and if severe enough results in failure of the system.

In Fig. 25 is shown the effect of frequency drift on signal bias in an *AM* system. In an *AM* system little bias is produced until the carrier reaches the cutoff region of the filter. The bias then becomes rapidly negative due to the increased loss and decreased amplitude of demodulated signal. When an automatic gain control arrangement is used the bias becomes positive due to a distorted envelope shape. The demodulated wave form determines the degree of sensitivity to frequency drift and depends on the bandwidth both before and after demodulation.

In an *FS* System using a linear discriminator, frequency drift changes the d-c. component of the signals and thus changes the operating point on the demodulated wave. The amount of bias depends upon the slope of the wave front and is thus affected by the amount of low-pass filtering. The effect of frequency drift on bias for a number of *FS* systems is shown in Figs. 26 and 27. If a two-bandpass filter type of discriminator is used the system

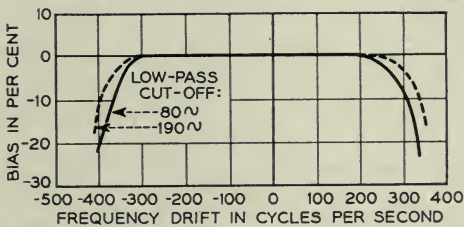


Fig. 25.—Signal bias versus frequency drift for *AM* transmission in a 740-cycle band.

is insensitive to moderate frequency drifts due to the flat pass bands as illustrated in Fig. 26. The relative shape and amplitude of the signal from a linear discriminator does not change appreciably with frequency drift; only a d-c. displacement occurs. This makes it desirable to have the low-pass filter coupled to the output amplifier by a network which passes only the useful signaling frequencies and blocks the d-c. and very slow drift components. When this is done the effect of frequency drift on bias is not greatly different from that for an *AM* system, as may be seen by comparing the dotted curves on Figs. 26 and 27 with Fig. 25.

The general method of performing this d-c. elimination is illustrated in the block diagram of Fig. 29. The output from the low-pass filter is passed through a coupling network which blocks the d-c. and passes the useful signaling frequencies. The output of the coupling network is passed through a positive feedback nonlinear amplifier which has but two output conditions representing the mark and space of the telegraph signal. The feedback network passes d-c. and low frequencies so as to just compensate for the loss of the coupling network. The time constant of the coupling network may be made large enough so that the signal wave form into the

d-c. amplifier is practically the same as at the output of the low-pass filter. The operating point on the demodulated wave may be readily adjusted by

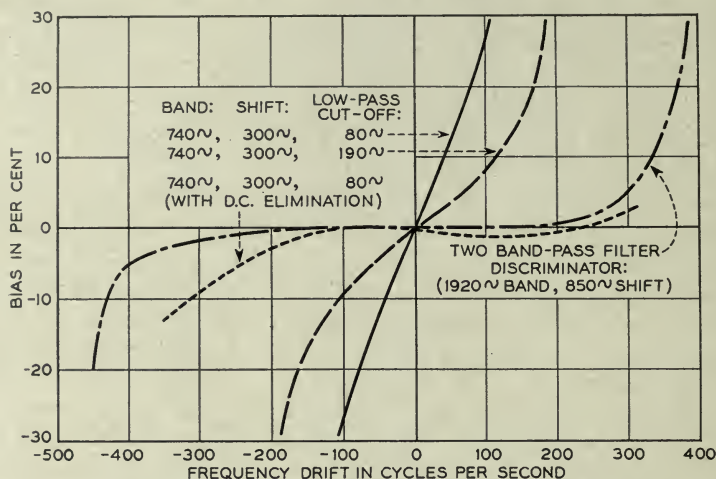


Fig. 26.—Signal bias versus frequency drift for FS transmission—wide filter bands.

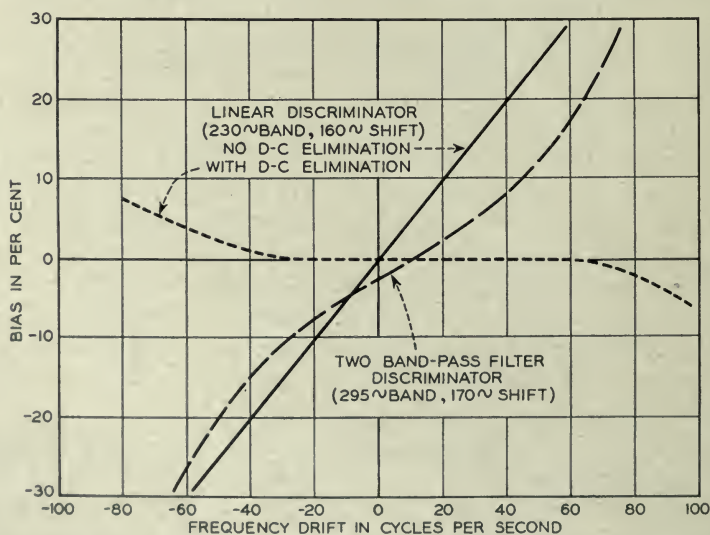


Fig. 27.—Signal bias versus frequency drift for FS transmission—narrow filter bands.

adjusting the bias voltage at the input to the d-c. amplifier. This arrangement differs from the impulse type of d-c. elimination, in which the demodulated wave form is effectively differentiated to form pulses but a fraction of a unit dot in length.

When d-c. elimination is used, operation with FS signals decentered in the pass band causes little bias but it does cause a loss in signal-to-noise ratio, especially at high noise levels. Due to the effect of noise peaks in causing a dip toward zero in the demodulator output, the effect of noise becomes exaggerated during the signal condition which is farther from midband. This change in signal-to-noise condition with frequency drift is shown in

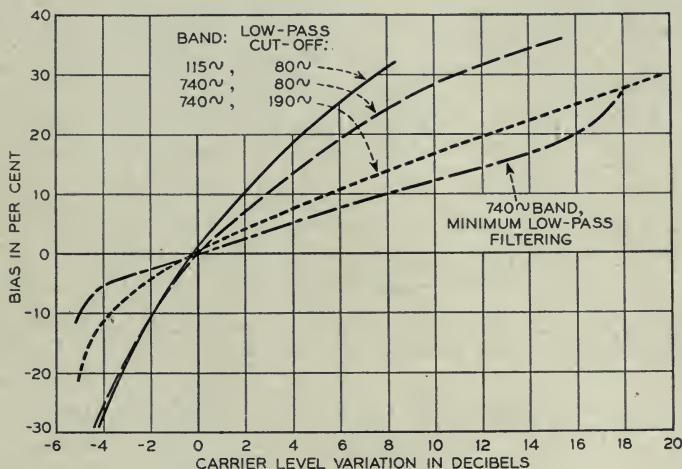


Fig. 28.—Signal bias versus carrier level variation for AM transmission.

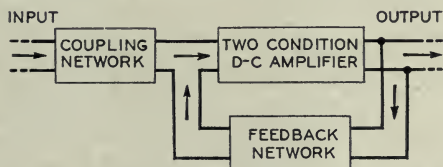


Fig. 29.—Block diagram of dc. elimination and restoration method used to minimize signal bias caused by frequency drift.

Figs. 24 and 30. A comparison between a two-bandpass-filter discriminator and a linear discriminator with d-c. elimination as regards frequency drifts in the presence of noise is shown in Fig. 31.

Radio telegraph systems operating in the H.F. region require a high degree of frequency stability if narrow bandwidths are to be used. Because of the several frequency conversions involved a number of different oscillators or frequency sources are involved but usually the major burden of frequency stability rests on the transmitter exciter and the high-frequency beating oscillator of the receiver.

Various methods of automatic frequency control may be used to hold the

carrier input to the demodulator at the correct frequency. In the case of FS signals the control may be arranged to operate only on the marking frequency or to utilize both the mark and space conditions. It is preferable to have inherent frequency stability rather than to compensate for the drift at the receiving end since it is difficult if not impossible to provide an automatic frequency control which will not reduce the transmission capabilities

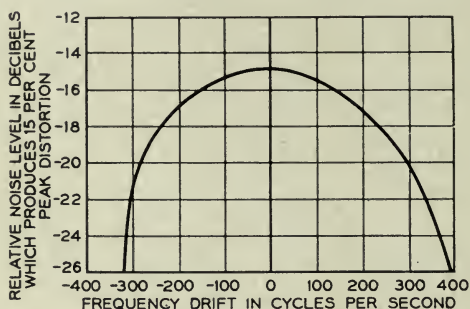


Fig. 30.—Effect of frequency drift on distortion produced by impulse noise in FS transmission—740-cycle band, 100 cycle frequency shift, 80-cycle cutoff low-pass filter.

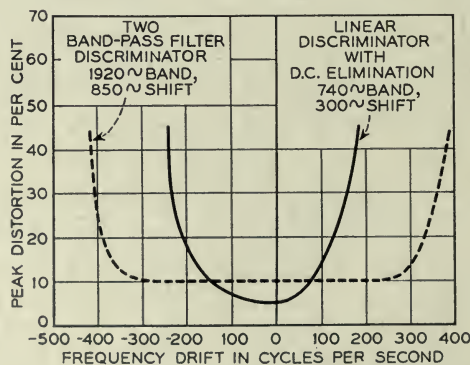


Fig. 31.—Linear discriminator versus two-bandpass-filter discriminator when frequency drift occurs in the presence of thermal noise—4 db rms noise-to-carrier ratio.

of the system in the presence of noise and other interference. Best results are obtained if the frequency stability is high enough to require but a very slow correction, which usually dictates some mechanical rather than electronic tuning arrangement. Manual retuning may be found satisfactory where stability is reasonably good provided care is taken in making the adjustments. Suitable frequency stability with the retention of flexibility in frequency adjustment may be obtained by frequency sources making use of a combination of crystal oscillators and high-stability variable oscillators of lower frequency.

Level Variations

Extreme and rapid variations of received level exist in H.F. radio transmission. It is upon the ability to accommodate these level variations that the merits of an H.F. radio telegraph system must largely be judged. FS telegraph shows its outstanding advantage in this respect.

In an *AM system* in order to obtain zero-bias signals and optimum signal-to-noise conditions the operating point must be near the half amplitude point on the demodulated wave. This means that complete failure will result for a drop in level of 6 db unless some compensating arrangement is provided. The slope of the bias-versus-level characteristic depends upon the slope of the demodulated wave which in turn depends on the bandwidth of the system and upon the degree of low-pass filtering. The bias-versus-level characteristics of some AM systems are shown in Fig. 28. Where the level variations are relatively slow compared to the signaling speed, automatic gain control circuits can be used to maintain a nearly constant level into the demodulator. However, where large rapid level changes occur, as in the H.F. range, it is seen that a narrow band AM system would fail completely regardless of the amount of transmitted power. For printer operation over an AM system in the H.F. range a fairly wide band and little low-pass filtering should be used, so as to keep the wave shape of the signals as square as possible and thus obtain a fairly flat bias-versus-level characteristic. By adjusting the operating point low on the demodulated wave, approaching the spacing noise level, the greatest possible range of acceptable rapid level change will be obtained. The slower level change components may be handled by the usual automatic gain-control circuits. This will cause the bias of the signals to average somewhat marking but the peak distortions will be kept to a minimum.

In an *FS system* no bias is produced so long as both the marking and spacing frequencies are affected alike, with their received levels remaining equal. Such non-selective fading conditions cause no distortion even when they occur at quite rapid rates. If a balanced type of discriminator is used, amplitude limiting is not essential to obtaining this immunity from non-selective variations in attenuation. It is only when the mark and space levels are different that bias results. In Fig. 32 are shown bias versus mark-to-space level ratio characteristics both with and without a limiter. More bias exists when there is no limiter because the amplitude of the demodulated wave is directly affected and consequently the low-pass filtering also becomes a factor. With a limiter the amplitude of the demodulated wave is held constant and the amount of low-pass filtering has no effect on bias. Some bias is still produced, however, due to the differently shaped frequency transients in the passband of the receiving system when a level change occurs

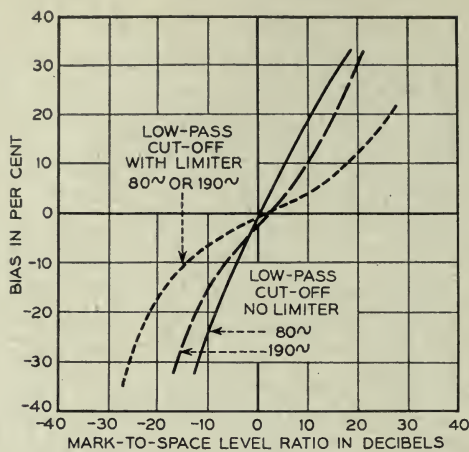


Fig. 32.—Signal bias versus mark-to-space level ratio in FS transmission—740-cycle band, 350-cycle frequency shift.

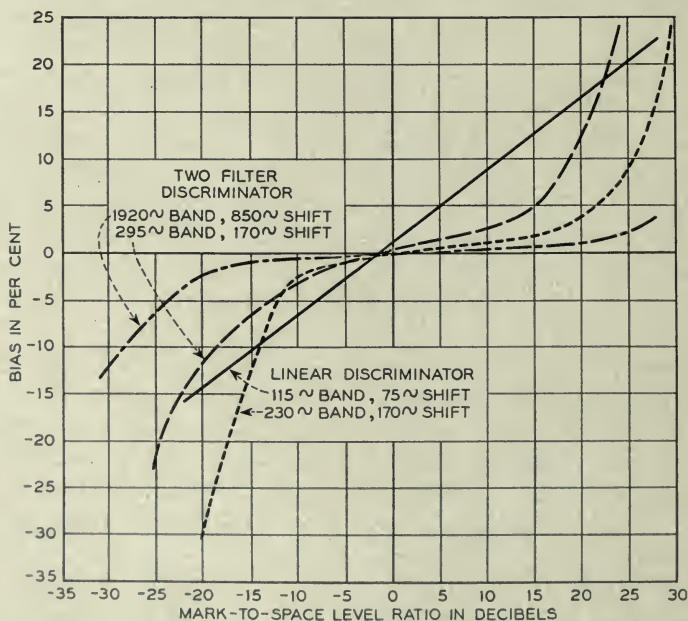


Fig. 33.—Signal bias versus mark-to-space level ratio in FS transmission—80-cycle cutoff low-pass filter.

at the moment when the frequency changes. In Fig. 33 this bias effect is demonstrated for various bandwidths. For moderate mark-to-space level ratios the bias effect is small and linear, with a slope which usually varies

inversely with a change in bandwidth. At extreme level differentials the bias may rise very rapidly due to the amplitude and phase characteristics of the input passband; transients from the greater amplitude condition may severely interfere with the lower amplitude condition. For the types of bandpass filters used in the tests it appeared that the amount of level difference required to produce 20% bias did not change greatly with bandwidth. Severe wave shaping of the signal at the transmitter was found to be an aid in reducing the bias effect due to such transients but the characteristic distortion became too great to give any practical improvement.

The fading modulator used in obtaining the data for Fig. 33 caused no change in phase. Selective fading over an actual radio circuit would involve considerable phase shift and greater distortion might be expected. The data of Fig. 32 were obtained by use of the phase control associated with the crystal filter of the radio receiver to vary the loss-versus-frequency characteristics of the receiving pass band and thus cause unequal mark and space amplitudes. This method gave an amplitude and phase characteristic for the transmission band more like that over an actual radio circuit.

MULTIPATH PROPAGATION EFFECTS

The rapid fading conditions prevailing in the H.F. range are brought about by multipath propagation. Under such conditions, the signal induced in a receiving antenna by a distant transmitter may be the resultant of two or three separate waves each propagated over a different path. If two waves arrive over paths differing in length by an odd number of half wavelengths the resulting 180° phase difference causes maximum cancellation. On the other hand if the paths differ in length by an integral multiple of whole wavelengths the waves arrive in-phase and maximum reinforcement results. The difference in path lengths may at times be as great as 500 to 1500 kilometers (delay times of 2 to 5 milliseconds) which in the H.F. region corresponds to thousands of wavelengths. Under these maximum conditions waves at one frequency may arrive in phase while waves at a frequency a few hundred cycles away may arrive in phase opposition. Since the path lengths are constantly changing, the transmission at a given frequency is subject to wide variations in amplitude and phase with time. When the difference in path lengths is not great enough to cause frequencies in one portion of a communication channel to fade differently from those in another portion the term "non-selective" or "flat" fading is applied. When the difference in path lengths becomes great enough to cause considerable amplitude or phase distortion over the transmission band the term "selective" fading is used. Since the propagation paths existing at a given moment vary for different antenna sites, the fading patterns obtained from two or three antennas separated by several wavelengths usually show a

considerable phase difference so that a given frequency is not likely to fade into the noise level at all antennas simultaneously. By employing separate receivers for each antenna and suitably combining or selecting the demodulated outputs, a system is obtained which is much less susceptible to fading. Such a method is called *space diversity* reception. Inasmuch as fading over a given combination of paths is highly selective with respect to frequency much the same effect is obtained by *frequency diversity* reception. When this method is employed the intelligence is transmitted on two or more frequencies simultaneously, and then received by separate receivers from a single antenna and the resulting demodulated signals combined or selected as for space diversity.

In telegraph transmission large differences in delay over two separate propagation paths cause the telegraph signal transitions to arrive at different instants over the two paths. Thus, there are intervals of overlap when a marking condition is received over one path and a spacing condition over a second path. When two components of nearly equal amplitude arrive at nearly 180° phase difference a signal transition may involve large and sudden amplitude and phase changes. The resulting transients in the bandpass networks of the receiving equipment may cause fortuitous distortions considerably greater than the difference in delay times over the two paths. The wider the pass band of the receiving system the shorter the duration of these fortuitous transients and hence the less the distortion. This phenomenon is one of the determining factors in the selection of bandwidth and frequency shift to be used in a given application of FS telegraphy. It becomes of increasing importance when the circuits are long and at higher signaling speeds such as are used in time-division multiplex methods.

In an AM system the effect of large differences in path lengths is usually a filling in of the spacing intervals with resulting marking bias. In an FS system the overlap time and associated transients may add to either marking or spacing intervals in a random fashion depending on the amplitude and phase conditions at each transition. The overlapping of the mark and space frequencies in FS transmission can sometimes be heard in an AM receiver as short pips of audio tone at each transition, the audio tone being the beat between the two frequencies.

Use of Superimposed Phase Modulation

Superimposed phase modulation has sometimes been employed as a simple means for achieving a certain amount of frequency diversity both in AM and FS telegraph systems. This consists in causing the radiated signal to oscillate continuously through a small phase angle at a rate relatively high compared to the dotting speed. Phase modulation spreads the energy of the signal over a wider frequency band so that the complete loss of the

signal through selective fading becomes less probable. The spectra generated by sinusoidal phase modulation of 1.0, 1.4, and 2.0 radians are shown in Fig. 34. Most of the energy is seen to be concentrated in the carrier and first order sidebands. Less than 1.0 radian of modulation results in too little amplitude of the sidebands, while more than 1.5 radians results in too wide spread of energy outside the first order sidebands. The center three

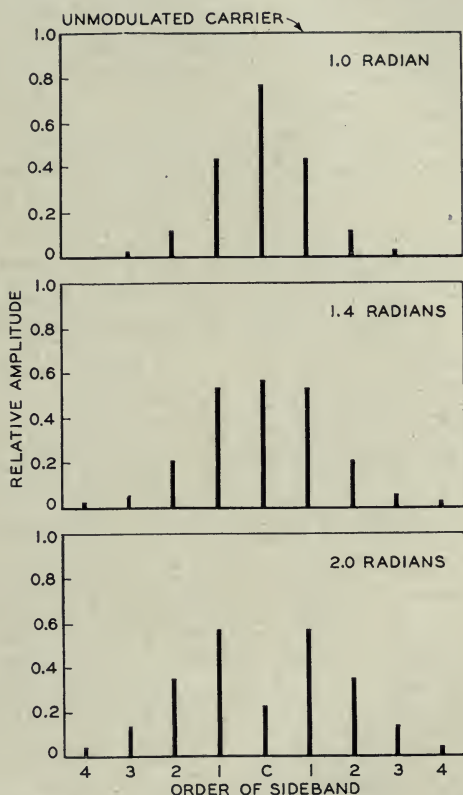


Fig. 34.—Frequency spectra for sinusoidal phase modulation.

components are equal at about 1.4 radians. In the case of FS the phase modulation frequency appears as a variation in amplitude of the signals from the discriminator. For AM no additional amplitude variation is caused by the phase modulation if there is no selective attenuation in the medium, but if such exists the phase modulation frequency or a multiple thereof appears in the rectified signal. To permit these unwanted amplitude variations to be removed by the low-pass filter so as not to break up the signals, the phase modulating frequency should preferably be ten or more times the maximum signaling frequency.

A test of superimposed phase modulation on FS signals was made over a radio circuit approximately 200 miles in length. A frequency shift of 850 cycles, and one radian of 200-cycle phase modulation, was used. A 60-word-per minute test sentence was transmitted and received without space diversity. It was found over a period of several hours that the phase modulation, on the average, gave a decrease in printed errors of about 50% when the error rate was in the proximity of 1 to 2%. For short intervals the reduction in errors was often considerably greater. The use of 200-cycle phase modulation when space diversity is used provides little or no improvement and is therefore undesirable.

A more effective way of employing phase modulation with FS signals would be to use a phase swing of ± 1.4 radians at a frequency of 2 to 3 times the frequency shift and to demodulate separately the three major components of the signal, thus obtaining in effect a triple-frequency diversity system. This of course involves quite a wide transmitted band, but it might be of use in cases where space diversity is impossible, such as on board ships. When a space diversity arrangement is feasible it is much to be preferred.

Diversity Operation

To obtain reliable operation in the H.F. range it is common practice to employ space diversity reception. The use of frequency diversity, with the increase of transmitted power and greater frequency space required, is seldom justified if space diversity reception can be arranged. For AM radio telegraph, double or triple-space diversity receiving arrangements are frequently used. Since an FS signal generally covers more frequency space, it is even more likely to be mutilated by selective fading than an AM signal. It has been found, however, that a double-space diversity system for FS signals usually gives sufficient diversity action provided it is of a type that permits switching between channels at signaling speed without causing appreciable distortion. This is necessary since it is a frequent occurrence that the mark of one channel may fade, leaving a good space, while the opposite may occur on the second channel. Since an FS system can accept rapid level changes, the main purpose of diversity methods is to insure that both the mark and space portions of the signal will be received above the noise level. In the case of AM telegraph, since it cannot accept rapid level changes, diversity operation is important not only in keeping the signal above the noise but also in averaging out some of the rapid level changes. For this reason AM systems usually show considerable improvement in going from double to triple diversity. It would be expected that a like change would show much less improvement in an FS system.

Diversity Channel Selection

The method employed to combine or select the channels of a diversity system is of great importance. For an AM system the relatively simple method of using a common load circuit for the diode detectors of the diversity channels is generally used. By deriving a common AVC voltage from the combined output and by properly adjusting the receiver sensitivities a fairly constant output is obtained. The parallel connection of the diode detectors causes the stronger signal to effectively block the weaker signal thus giving a fairly sharp diversity selection characteristic. The problem of combining the diversity channels of an FS system is more complicated mainly because of the amplitude limiting. If amplitude limiting is used in each diversity channel before demodulation, the resulting constant amplitude signals convey no information as to their relative amplitudes as received from the antennas. Any diversity selection must then be obtained by some indirect method. It is necessary to furnish some selecting device since the noise from a faded channel, if added directly to a good signal from another channel, will cause high distortion.

In an early frequency shift system employing a two-bandpass filter discriminator (shown previously in Fig. 17) it was found that for a poor signal-to-noise condition the sum of the outputs of the mark and space rectifiers increased above that for a good signal-to-noise condition. This increase was utilized to suppress the output of the poorer channel and emphasize that of the better channel. Although neither the degree of diversity selection nor the speed of response was as good as might be desired, fairly satisfactory results were obtained.

Another method which has been used involves the derivation of control currents or voltages proportional to the amplitudes of the incoming signals which in turn select the better diversity channel by some type of gate action. The time constants of the control circuits must be low enough to permit switching at signaling speed without introducing considerable distortion. The gate circuits must also be of a type which does not introduce interfering transients or otherwise allow the control voltages or currents to interfere with the signal. This method permits very sharp diversity selection and has the capability of approximating ideal results although it becomes somewhat involved in a practical form.

A considerably simpler method has been used in some of the more recent FS terminals. It is based on the use of a single-amplitude-limiter through which pass the signals of both diversity channels. This is made possible by arranging the two signals at the input to the limiter to be at different frequencies. At the output of the limiter the two signals are separately demodulated and then combined. When one of the signals is considerably

greater in amplitude than the other at the input to the limiter the relative difference in level is increased by an additional amount of about 6 db at the limiter output. The limiter output may be considered as the stronger signal frequency-modulated by the weaker signal. For small modulation indices the amplitude of the first order sideband is approximately one-half the modulation index thus explaining the 6 db added difference in level at the limiter output. As the input levels approach equality the added level difference decreases to zero. A block diagram indicating the arrangement of such a diversity system is shown in Fig. 35. Tests were made of both parallel and series connections of the two discriminator outputs. With a parallel connection the discriminator having the greater output blocks the rectifier output of the other discriminator and thus gives a sharp diversity selection characteristic. However, the level ratio of the channels at which a switch

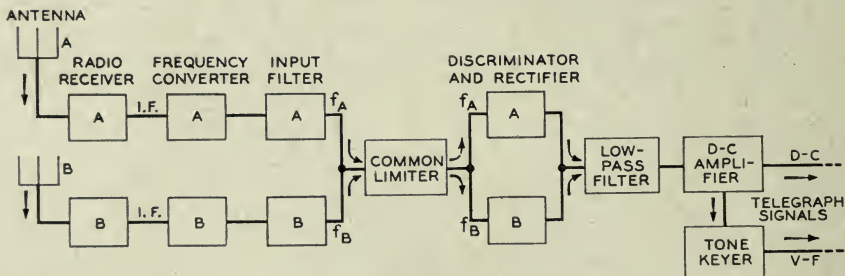


Fig. 35.—Block diagram of dual-diversity FS receiving system using a common limiter.

takes place is affected not only by the ratio at the limiter input but also by frequency drift and discriminator slope. With the series connection only the input level ratio at the limiter input affects the diversity switching; this arrangement was therefore selected as the preferred method although its selection characteristic is not sharp. The series and parallel combining characteristics are shown in Figs. 36 and 37.

Various tests were made on a terminal having a 1500-cycle bandwidth and using the series combining method to determine the signal-to-noise characteristics under different conditions of diversity fading. A frequency shift of 850 cycles was used and the midband frequencies of the two diversity channels at the common limiter input were 30 and 35 Kc. Figure 38 shows the distortion versus signal-to-noise ratio characteristics of each channel separately and in diversity combination for various relative level conditions of the two channels. During diversity operation equal noise levels were maintained in the two channels and various combinations of level differences of the two channels were preserved as the whole signal level combination was varied. The level differentials are indicated in

the figure for each curve. The signal-to-noise level scales refer to the highest level portion of the diversity signal. Since the amplitude modulator which was used to simulate the selective fading did not produce phase shifts or

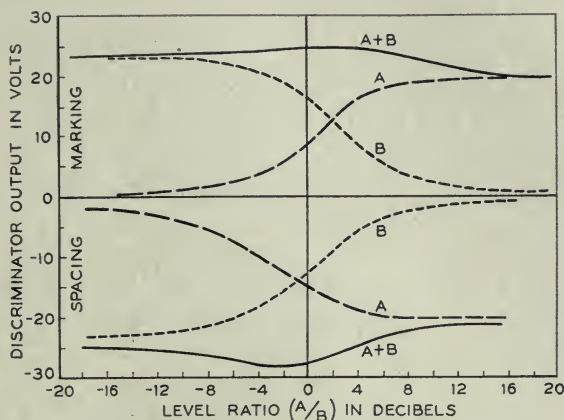


Fig. 36.—Diversity combination characteristic obtained by series addition of discriminator outputs—levels measured at output of 400 kc I.F. amplifier.

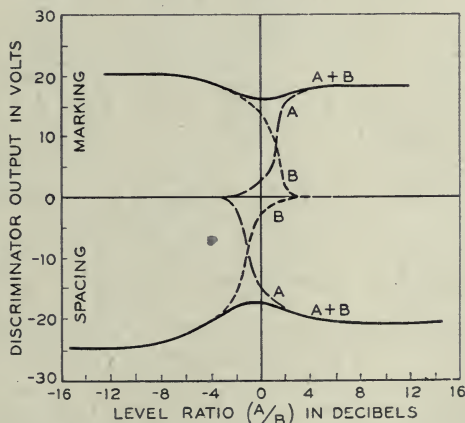


Fig. 37.—Diversity combination characteristic obtained by parallel addition of discriminator outputs—levels measured at output of 400 kc I.F. amplifier.

transmission delays as would actually exist over H.F. radio circuits the actual distortions shown by the curves are optimistic.

The ideal diversity selection circuit should theoretically give a signal-to-noise characteristic identical to that of a single channel under the signal-to-noise condition corresponding to the signal of the best momentary reception. It will be seen that the test results of Fig. 38 approach this limit within 2 or 3 db at a peak distortion of 20%. Part of this difference is

due to the dissimilar bandpass characteristics of the two channels of the experimental unit used for the tests and part due to the lack of an extremely sharp diversity selection. It should be pointed out that the conditions under which the theoretical maximum diversity signal-to-noise condition may be reached are very hard to obtain in practice. If the noise levels in the two

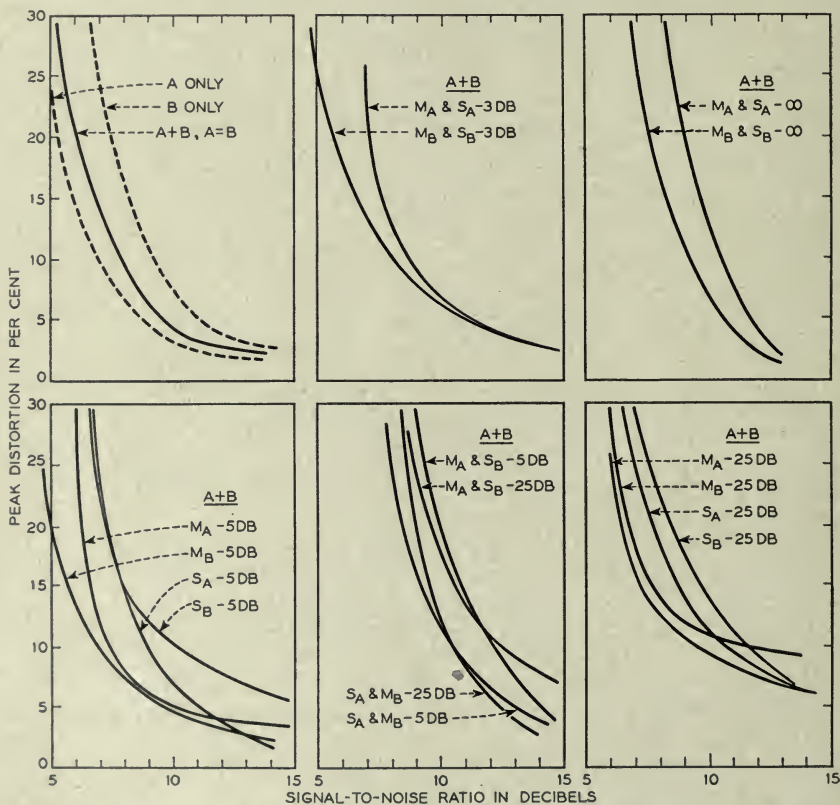


Fig. 38.—Peak distortion versus signal-to-noise ratio characteristics for dual-diversity operation with a common limiter. Signal-to-noise measured at output of 400 kc I.F. amplifier.

channels are not equal, and if the diversity selecting method does not at all times exclusively select the channel with the greater signal, the distortion characteristic will deteriorate accordingly. Because the AVC sensitivities of two radio receivers may differ considerably the noise levels cannot be maintained closely the same and usually no provision is made for determining the noise level except by ear. The slightly better diversity action which can theoretically be obtained is therefore felt to be of doubtful usefulness under actual operating conditions. The common limiter method has the

advantage of being simple in that the diversity action is obtained in the transmission circuits directly and no added switching circuits or adjustments are required. Tests of this type of diversity selection in the field have indicated a marked superiority over earlier FS terminal equipment.

In connection with the use of a common limiter care must be taken in the selection of the two-channel frequencies. Frequencies having nearly integral ratios such as 3:5 and 5:7 produce disturbing amplitude modulation of the demodulated signal. The frequencies should be chosen so as to avoid low integral ratios; then all amplitude modulations are negligible or easily filtered out. Where frequency drift is to be allowed for, the frequencies should be chosen so as not to approach a low integral ratio at any place in the expected drift range.

If the radio receivers associated with a space diversity FS system have automatic gain control it must be a common control so the receivers will change gain equally. The use of common AVC prevents overloading of the receivers as the received signal strength varies. If no common AVC is available the receivers should be operated in the manual gain-control condition.

CONCLUSIONS

General Comparison of FS and AM Carrier Telegraphy

The foregoing sections have compared the characteristics of AM and FS carrier telegraph transmission under various conditions. Whether or not FS would prove to be the preferable method for a specific communication use depends largely on the transmission medium and the quality of transmission desired. As regards frequency space requirements, both methods provide essentially the same signaling speed capability for a given bandwidth.

As to the ability to transmit through noise, FS has an advantage of 3 to 4 db at distortions approaching the failure point when equal bandwidths are compared. At lower distortions the advantage of FS is 6 db or more so that it is attractive in this respect for tandem operation of several telegraph sections where regeneration of signals is not practiced. When frequency space permits wider bands, with correspondingly increased frequency shifts, the signal-to-noise advantage of FS over AM increases for low noise levels. Wide band FS therefore provides a means of obtaining higher quality circuits if the noise level is not too great.

The AM method is basically less susceptible to frequency variations than is the FS method. However, as has been illustrated, frequency drift can be compensated for by d-c. elimination so as to make FS comparable to AM in this respect.

FS transmission is essentially immune from effects of non-selective level variation, even when extremely rapid, and in this characteristic displays its most outstanding advantage over AM.

Operation Over Wire Circuits

A wire circuit usually provides a transmission medium having a low noise level with slow and relatively small variations in attenuation. Such circuits, when equipped with suitable automatic gain control, allow stable operation with AM telegraphy and but little improvement could probably be obtained by using FS. The choice between AM and FS under such relatively ideal conditions becomes one of economic considerations of the terminal equipment and carrier supply. However, when FS is applied to multichannel systems the problem of interchannel interference requires attention. For wire circuits having high noise levels or sudden changes in attenuation the use of FS instead of AM provides considerable improvement and in severe cases the FS method may be a necessity for satisfactory operation. Wide band FS operation with its sharper breaking distortion-versus-noise-level characteristic gives a low value of rms-to-peak distortion which would be especially advantageous for tandem operation. However, the necessary frequency space for wide-band operation is not usually economically justified for wire line operation.

Operation Over Radio Circuits

For operation over radio circuits providing stable conditions similar to those on wire circuits the FS method does not show a great advantage over the AM method. In the case of long distance telegraphy in the H.F. range, however, FS shows a marked advantage over AM. This is because of the rapid fading and high noise conditions which commonly prevail in the H.F. region. The amount of rapid variation in marking level that an AM system can accommodate is less than the difference between marking and spacing levels that an FS system can tolerate. In the worst case of selective fading the level differences between the mark and space frequencies might approach values equal to the short time level swings of a single frequency, but in general would be less. A given condition of selective fading thus causes less distortion in an FS system than in an AM system. FS allows the use of narrow bands without much loss in signal quality in the presence of fading, whereas AM does not. FS therefore is essential for satisfactory operation of closely spaced narrow band H.F. radio channels. Where frequency space is not restricted and wider bands are used to permit considerable frequency drift, the improvement afforded by FS over AM is materially less. To obtain optimum results from an AM system, however, requires

more careful adjustment and more attention than does an FS system. This is partly due to the amplitude limiter in the FS system which results in a constant amplitude of signal from the discriminator and partly due to the fact that an FS signal is no more subject to noise interference during the spacing condition than during the marking condition. Therefore the operating point on the demodulated wave may be set and left for long intervals even though transmission conditions vary widely. This greater ease in maintaining good adjustment of the equipment probably accounts for some of the apparent improvement in changing from an AM to an FS system.

It should be noted that a system may fail either because of level variations well above the noise level or because of the signal becoming submerged in noise. If a system fails because it can accept only moderate level variations, an increase in transmitted power will provide no improvement since the level variations will remain the same as before. On the other hand, a system which can accept very wide variations in level will show improvement upon increasing transmitted power up to the point where no failures occur due to an unfavorable signal-to-noise ratio.

The over-all improvement obtained in changing from AM to FS radio telegraph is sometimes expressed as a ratio of transmitted powers required to give equivalent transmission results over the two systems. Such a ratio fluctuates widely depending upon the prevailing conditions. With little fading the improvement ratio will be mainly due to the better signal-to-noise obtained with FS and may be less than 5 db. Under severe fading conditions no amount of power may give good results with AM while FS may be satisfactory.* Thus the power ratio would become infinite. By making a long-time comparison an average power ratio figure may be found which gives equal average error rates in the printed copy from each system. Such tests⁷ between a triple space diversity AM system and a double space diversity FS system have indicated a power ratio of 11 db in favor of the latter when the error rate was 0.1 to 0.5 per cent.

When the two systems are thus made equal by adjustment of transmitted power, more errors due to the signal becoming submerged in noise occur in the FS system to compensate for a larger number of errors in the AM system due to rapid level changes. Often the reason for changing a radio telegraph system from AM to FS is to increase the reliability of the circuit and not just to save transmitted power. To insure a definite improvement in such cases the carrier level should not be decreased more than about 6 db.

REFERENCES

1. Certain Topics in Telegraph Transmission Theory, H. Nyquist, *Trans. of A.I.E.E.*, Vol. 47, April 1928, pp. 617-644.
2. Frequency Modulation, Balh. van der Pol, *Proc. of I.R.E.*, Vol. 18, July 1930, p. 1202.

3. Measurement of Telegraph Transmission, H. Nyquist, R. B. Shanck, and S. I. Cory, *Trans. of A.I.E.E.*, Vol. 46, Feb. 1927, pp. 367-376.
4. Frequency Modulation Noise Characteristics, M. G. Crosby, *Proc. of I.R.E.*, Vol. 25, April 1937, p. 472-514.
5. Performance Characteristics of Various Carrier Telegraph Methods, T. A. Jones, K. W. Pfleger, *Bell Sys. Tech. Jour.*, Vol. 25, pp. 483-531, July 1946.
6. Transients in Frequency Modulation, H. Salinger, *Proc. of I.R.E.*, August 1942, pp. 378-383.
7. Observations and Comparisons on Radio Telegraph Signaling by Frequency Shift and On-Off Keying, H. O. Peterson, J. B. Attwood, H. E. Goldstine, G. E. Hansell, R. E. Schock, *RCA Review*, March 1946.

Reflections from Circular Bends in Rectangular Wave Guides—Matrix Theory

By S. O. RICE

A method of computing reflections produced by circular bends in rectangular wave guides is presented. The procedure employs the theory of matrices. Although the matrix equations are quite simple, a considerable amount of calculation is necessary before quantitative results may be obtained. Fortunately, the approximate formulas pertaining to gentle bends hold surprisingly well for rather sharp bends. These formulas are obtained by a limiting process from the matrix equations. The approximate formula for reflection from an H-bend (in which the magnetic vector lies in the plane of the bend) generalizes an earlier result due to R. E. Marshak. The corresponding formula for the E-bend appears to be new.

INTRODUCTION

A NUMBER of investigators have studied the propagation of electromagnetic waves in a bent pipe of rectangular cross-section, the bend being along an arc of a circle. H. Buchholz¹, S. Morimoto², and W. J. Albersheim³ have employed Bessel functions to express the field in the bend. The form assumed by the field when the radius of curvature of the bend becomes large has been obtained by K. Riess⁴ and R. E. Marshak⁵ who use approximations suited to this case. Marshak also obtains expressions for various reflection and transmission coefficients. A discussion of the subject using rather simple but approximate analysis is given on pages 324–330 of a text book⁶ by S. A. Schelkunoff. The Bessel function approach is also sketched in the same section.

Here we study the disturbance produced when a wave goes around a circular bend (of some given angle) in a rectangular wave guide, the guide being straight on either side of the bend. Especial attention is paid to the dominant mode reflection coefficients g_{10}^- and d_{01}^- corresponding to H-bends and E-bends, respectively. As equations (4.2–6) and (4.4–4) show, these reflection coefficients (which are of the nature of voltage rather than power reflection coefficients) vary inversely as the square of the radius of curvature of the bend when the bend is gentle. The substance of (4.2–6) has been given by Marshak⁵ for the important case in which only the dominant mode is propagated and the angle of the bend not too small.

When the bend is so sharp that the formulas mentioned above do not apply the reflection coefficients may be computed from the rather simple looking matrix expressions (2.3–3) together with (2.3–4). However, their appearance is deceptive and, as is shown by the numerical work in Part V, considerable labor is necessary to obtain an answer.

The gentle bend formulas were obtained from the matrix equations by the limiting process described in Part III. It seems likely that the matrix method, which is similar to the method used in an earlier paper⁷ on transmission line equations, may be applied to other wave guide problems. With this thought in mind, the development of Parts II and III has been couched in general terms.

The matrices used in the present theory are of infinite order since the guide may support an infinite number of modes of propagation. This fact makes it difficult to justify all the steps in our analysis, and we do not attempt to do so.* Despite this lack of rigor, I believe that the procedures given here lead to the correct results since they yield, for gentle bends, expressions obtained by Buchholz and Marshak. Moreover, although numerical results tabulated in Part V were obtained by using matrices of only the second and third order, they indicate a rapid convergence as the matrix order is increased.

PART I

PROPAGATION OF WAVES IN GUIDE

1.1 *Propagation in a Straight Wave Guide*

Rather general expressions for the electric and magnetic intensities E and H in a field are (see pp. 127-128 of Reference⁶)

$$\begin{aligned} E &= -i\omega\mu\vec{A} + \frac{1}{i\omega\epsilon} \text{grad div } \vec{A} - \text{curl } \vec{B} \\ H &= \text{curl } \vec{A} + \frac{1}{i\omega\mu} \text{grad div } \vec{B} - i\omega\epsilon\vec{B} \end{aligned} \quad (1.1-1)$$

The field is assumed to vary with the time t as $e^{i\omega t}$, ω is the radian frequency, μ the permeability and ϵ the dielectric constant (for free space $\mu = 1.257 \times 10^{-6}$ henries/meter, $\epsilon = 8.854 \times 10^{-12}$ farads/meter). The vector potentials $\vec{A}$ and $\vec{B}$ satisfy the wave equations

$$\begin{aligned} \nabla^2 \vec{A} &= \sigma^2 \vec{A}, & \nabla^2 \vec{B} &= \sigma^2 \vec{B} \\ \nabla^2 &\equiv \text{Laplacian operator} \\ \sigma^2 &= \omega^2 \mu \epsilon \end{aligned} \quad (1.1-2)$$

In dealing with bends, it is convenient to choose $\vec{A}$ and $\vec{B}$ normal to the plane of the bend. In our notation, this plane is always taken to be the x, z plane so that $\vec{A}$ and $\vec{B}$ are parallel to the y axis. The z axis is parallel

* Similar questions arise in the rigorous treatment of an infinite set of linear equations. A discussion of this subject is given in Chap. III of Reference⁸.

to the guide axis and, for the straight guide of the section, the guide walls are sections of the planes $x = 0, x = a, y = 0, y = b$.

Thus, a general wave traveling in the positive z direction may be described by the two functions (which represent the magnitudes of $\vec{A}$ and $\vec{B}$)

$$A = \sum_{m,n} g_{mn}^+ e^{-\Gamma_{mn}z} \sin(\pi mx/a) \cos(\pi ny/b) \quad (1.1-3)$$

$$m = 1, 2, 3, \dots; \quad n = 0, 1, 2, \dots$$

$$B = \sum_{m,n} d_{mn}^+ e^{-\Gamma_{mn}z} \cos(\pi mx/a) \sin(\pi ny/b) \quad (1.1-4)$$

$$m = 0, 1, 2, \dots; \quad n = 1, 2, 3, \dots$$

where the coefficients g_{mn}^+ and d_{mn}^+ are constants and the plus signs indicate propagation in the positive z direction.

The propagation constant Γ_{mn} is obtained from

$$\Gamma_{mn}^2 = \sigma^2 + (\pi m/a)^2 + (\pi n/b)^2, \quad \sigma = i2\pi/\lambda_0, \quad (1.1-5)$$

$\lambda_0 = \text{wavelength in free space.}$

Equation (1.1-5) arises when the typical term in (1.1-3) is substituted for A in the equation

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2} = \sigma^2 A \quad (1.1-6)$$

This and a similar equation for B are the forms assumed by (1.1-2) for the rectangular coordinates of our straight guide.

The electric and magnetic intensities in the guide are given by

$$\begin{aligned} E_x &= \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial x \partial y} + \frac{\partial B}{\partial z} & H_x &= -\frac{\partial A}{\partial z} + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial x \partial y} \\ E_y &= -i\omega\mu A + \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial y^2} & H_y &= -i\omega\epsilon B + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial y^2} \\ E_z &= \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial z \partial y} - \frac{\partial B}{\partial x} & H_z &= \frac{\partial A}{\partial x} + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial z \partial y} \end{aligned} \quad (1.1-7)$$

which follow from (1.1-1).

It is seen that the wave is completely specified by the g_{mn}^+ 's and d_{mn}^+ 's. These may be arranged as (infinite) column matrices in any convenient order. Thus in dealing with (1.1-3) and (1.1-4) we may write

$$g^+ = \begin{bmatrix} g_{10}^+ \\ g_{20}^+ \\ g_{11}^+ \\ g_{30}^+ \\ g_{21}^+ \\ g_{12}^+ \end{bmatrix} \quad d^+ = \begin{bmatrix} d_{01}^+ \\ d_{02}^+ \\ d_{11}^+ \\ d_{03}^+ \\ \vdots \\ \vdots \end{bmatrix} \quad (1.1-8)$$

In our work we shall consider only those modes corresponding to a fixed value of m (or of n) and the order is almost automatically fixed.

The factors which determine the propagation of the typical terms in the summations (1.1-3) and (1.1-4) for A and B are

$$\alpha_{mn}(z) = g_{mn}^+ e^{-z\Gamma_{mn}}, \quad \beta_{mn}(z) = d_{mn}^+ e^{-z\Gamma_{mn}} \quad (1.1-9)$$

The column matrices obtained by arranging these quantities in the same order as in (1.1-8) will be denoted by $\alpha(z)$ and $\beta(z)$. We may write

$$\alpha(z) = e^{-z\Gamma_\alpha} g^+, \quad \beta(z) = e^{-z\Gamma_\beta} d^+ \quad (1.1-10)$$

where $\exp(-z\Gamma_\alpha)$ and $\exp(-z\Gamma_\beta)$ are square matrices defined by power series each term of which is a square matrix:

$$e^{-z\Gamma} = I - \frac{z\Gamma}{1!} + \frac{z^2\Gamma^2}{2!} - \frac{z^3\Gamma^3}{3!} + \dots \quad (1.1-11)$$

I is the unit matrix and Γ_α is the diagonal matrix*

$$\Gamma_\alpha = \begin{bmatrix} \Gamma_{10} & 0 & 0 & \cdot \\ 0 & \Gamma_{20} & 0 & \cdot \\ 0 & 0 & \Gamma_{11} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad (1.1-12)$$

in which the order of the diagonal elements is the same as the order of the elements in the column matrix g^+ . Similarly Γ_β is a diagonal matrix whose elements are $\Gamma_{01}, \Gamma_{02}, \Gamma_{11}, \Gamma_{03}, \dots$, the order being fixed by d^+ . When Γ is replaced by Γ_α in (1.1-11) it is easy to obtain $\Gamma_\alpha^2, \Gamma_\alpha^3$, etc. and sum the resulting series to obtain

$$e^{-z\Gamma_\alpha} = \begin{bmatrix} e^{-z\Gamma_{10}} & 0 & 0 & \cdot \\ 0 & e^{-z\Gamma_{20}} & 0 & \cdot \\ 0 & 0 & e^{-z\Gamma_{11}} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad (1.1-13)$$

A similar expression exists for $\exp(-z\Gamma_\beta)$. The expression (1.1-10) for $\alpha(z)$ is seen to be true when the square matrix (1.1-13) is multiplied, by matrix multiplication, into the column g^+ .

It turns out that the field in a circular bend (in a rectangular guide) may be represented by a generalization of the foregoing expressions. In this generalization, which will be studied in the following sections, the square matrices Γ_α and Γ_β no longer have the simple form of diagonal matrices.

* That is, a square matrix in which all of the elements other than those in the principal diagonal are zero.

1.2 Propagation in a Circular Bend

In dealing with a circular bend we choose cylindrical coordinates (ρ, φ, y) as shown in Fig. 1. With these coordinates we associate new coordinates, shown in Figs. 1 and 2, (x, y, z) which have approximately the same significance as in the straight guide. z is the distance measured along the axis of

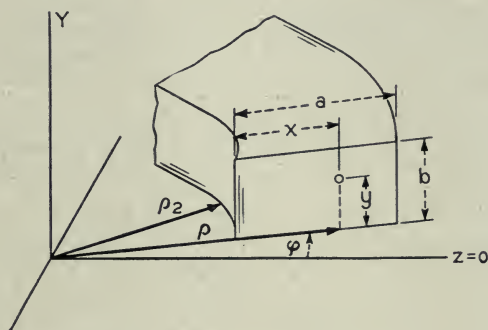


Fig. 1

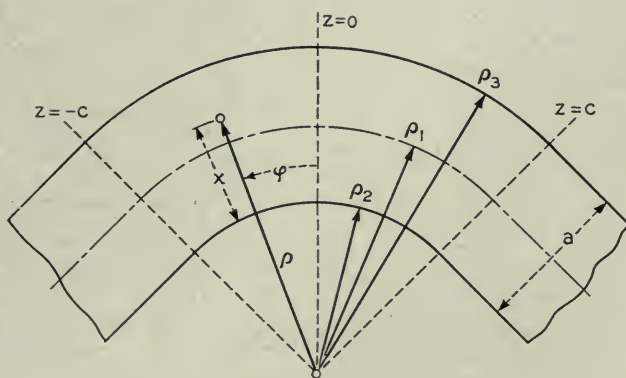


Fig. 2

the guide (defined as the locus of the centers of gravity of the transverse cross-sections of the guide), and x and y are the transverse coordinates.

Let $\rho = \rho_1 = (\rho_2 + \rho_3)/2 = \rho_2 + a/2$ be the radius of curvature of the guide axis, and let the origin of the polar coordinates be taken at the center of curvature. Then z is equal to $-\rho_1\varphi$ where the minus sign is necessary to make (x, y, z) a right-handed coordinate system. Since the vertical (in

Fig. 1) walls are to be specified by $x = 0$ and $x = a$ we set $x = \rho - \rho_1 + a/2$. Thus, the two sets of coordinates are related by

$$\begin{aligned}\rho &= x + \rho_1 - a/2 = x + \rho_2 \\ \varphi &= -z/\rho_1 \\ y &= y\end{aligned}\tag{1.2-1}$$

where ρ_1 , ρ_2 and a are constants.

We again choose $\vec{A}$ and $\vec{B}$ in (1.1-1) to be parallel to the y axis. In the cylindrical coordinates,

$$\begin{aligned}E_\rho &= \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial \rho \partial y} - \frac{1}{\rho} \frac{\partial B}{\partial \varphi} & H_\rho &= \frac{1}{\rho} \frac{\partial A}{\partial \varphi} + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial \rho \partial y} \\ E_\varphi &= \frac{1}{i\omega\epsilon\rho} \frac{\partial^2 A}{\partial \varphi \partial y} + \frac{\partial B}{\partial \rho} & H_\varphi &= -\frac{\partial A}{\partial \rho} + \frac{1}{i\omega\mu\rho} \frac{\partial^2 B}{\partial \varphi \partial y} \\ E_y &= -i\omega\mu A + \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial y^2} & H_y &= -i\omega\epsilon B + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial y^2}\end{aligned}\tag{1.2-2}$$

where now, from (1.1-2), A satisfies the wave equation

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left[\rho \frac{\partial A}{\partial \rho} \right] + \frac{1}{\rho^2} \frac{\partial^2 A}{\partial \varphi^2} + \frac{\partial^2 A}{\partial y^2} = \sigma^2 A\tag{1.2-3}$$

and likewise for B .

One method of dealing with (1.2-3) which is sometimes used is to assume

$$A = e^{ip\varphi} \times (\text{sine or cosine function of } y) \times f(\rho)\tag{1.2-4}$$

where $f(\rho)$ turns out to be a Bessel function of order p with its argument proportional to ρ . However, we shall proceed in a different direction.

The change of coordinates (1.2-1) transforms (1.2-2) into

$$\begin{aligned}E_x = E_\rho &= \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial x \partial y} + \frac{\rho_1}{\rho} \frac{\partial B}{\partial z} & H_x = H_\rho &= -\frac{\rho_1}{\rho} \frac{\partial A}{\partial z} + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial x \partial y} \\ E_y &= -i\omega\mu A + \frac{1}{i\omega\epsilon} \frac{\partial^2 A}{\partial y^2} & H_y &= -i\omega\epsilon B + \frac{1}{i\omega\mu} \frac{\partial^2 B}{\partial y^2} \\ E_z = -E_\varphi &= \frac{\rho_1}{i\omega\epsilon\rho} \frac{\partial^2 A}{\partial z \partial y} - \frac{\partial B}{\partial x} & H_z = -H_\varphi &= \frac{\partial A}{\partial x} + \frac{\rho_1}{i\omega\mu\rho} \frac{\partial^2 B}{\partial z \partial y}\end{aligned}\tag{1.2-5}$$

and (1.2-3) into

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\rho_1^2}{\rho^2} \frac{\partial^2 A}{\partial z^2} + \frac{1}{\rho} \frac{\partial A}{\partial x} - \sigma^2 A = 0,\tag{1.2-6}$$

where ρ_1 is a constant and $\rho = x + \rho_1 - a/2$ is to be considered a function of x . To solve (1.2-6) and the corresponding equation for B we assume

$$A = \sum_{m,n} \alpha_{mn}(z) \sin(\pi mx/a) \cos(\pi ny/b) \quad (1.2-7)$$

$$m = 1, 2, 3, \dots; \quad n = 0, 1, 2, \dots$$

$$B = \sum_{m,n} \beta_{mn}(z) \cos(\pi mx/a) \sin(\pi ny/b) \quad (1.2-8)$$

$$m = 0, 1, 2, \dots; \quad n = 1, 2, 3, \dots$$

these expressions being suggested by (1.1-3) and (1.1-4). The expressions (1.2-5) for the electric intensity show that this choice of A and B make its tangential component vanish at the walls of the guide. Thus the boundary conditions are satisfied.

In order to determine $\alpha_{mn}(z)$ so that the differential equation for A is satisfied, we substitute (1.2-7) in (1.2-6). The resulting left hand side of (1.2-6) may be regarded as a function, say $f(x, y)$, of x and y with the α 's and their derivatives entering as parameters. We must choose the α 's so as to make this function zero. Relations which must be satisfied by the α 's may be obtained by expanding $f(x, y)$ in a double Fourier series for which the typical term is a coefficient times $\sin(\pi mx/a) \cos(\pi ny/b)$, and then setting the coefficient of each term to zero. This form of expansion is suggested by (1.2-7). However, it should be mentioned that such an expansion is best suited to a function which vanishes at $x = 0$ and $x = a$, a condition not fulfilled by $f(x, y)$ because of the term $\rho^{-1} \partial A / \partial x$ in (1.2-6). This causes no real trouble because our region of representation runs only from $x = 0$ to $x = a$ and hence our series is no worse than the Fourier sine series for the periodic function (of period $2a$) which is -1 for $-a < x < 0$ and $+1$ for $0 < x < a$.

To carry out the procedure outlined above, we multiply (1.2-6) (after putting in (1.2-7)) by $\sin(\pi px/a) \cos(\pi \ell y/b)$ and integrate x from 0 to a and y from 0 to b . Using the expression (1.1-5) for Γ_{mn}^2 and reducing gives

$$-\Gamma_{p\ell}^2 \alpha_{p\ell}(z) + \sum_{m=1}^{\infty} [P_{pm} \alpha_{m\ell}''(z) - S_{pm} \alpha_{m\ell}(z)] = 0 \quad (1.2-9)$$

where p may have any one of the values $1, 2, 3, \dots$ and the double prime on α denotes the second derivative with respect to z . The P 's and S 's are constants given by

$$P_{pm} = (2/a) \int_0^a (\rho_1^2/\rho^2) \sin(\pi px/a) \sin(\pi mx/a) dx, \quad (1.2-10)$$

$$S_{pm} = -2\pi ma^{-2} \int_0^a \sin(\pi px/a) \cos(\pi mx/a) dx/\rho \quad (1.2-11)$$

The evaluation of these integrals is discussed in Appendix I. Thus (1.2-9) is the p^{th} equation of a set of differential equations to be solved simultaneously for $\alpha_{1\ell}(z)$, $\alpha_{2\ell}(z)$, $\dots$.

The customary method of solving a set of equations such as (1.2-9) is to assume that all the α 's vary as $e^{\gamma z}$ so that for each $\alpha_{m\ell}(z)$ we may write $e^{\gamma z} g_{m\ell}$. This leads to a set of simultaneous homogeneous linear equations for the constants $g_{m\ell}$. In order that these equations may have a solution the determinant of the coefficients must vanish. Since the only derivative of $\alpha_{m\ell}(z)$ contained in (1.2-9) is the second, γ appears in the determinant only as γ^2 . Let $\gamma_1^2, \gamma_2^2, \gamma_3^2, \dots$ be the values of γ^2 which cause the determinant to vanish and let $k_{1j}, k_{2j}, \dots$ be the values of $g_{1\ell}, g_{2\ell}, \dots$ corresponding to $\gamma^2 = \gamma_j^2$. The k 's are determined to within an arbitrary multiplying constant which, for the sake of convenience, is chosen so that $k_{jj} = 1$.

Thus one solution of the differential equation (1.2-6) is

$$A = e^{-\gamma_j z} \cos(\pi \ell y/b) \sum_{m=1}^{\infty} k_{mj} \sin(\pi m x/a). \quad (1.2-13)$$

This particular solution corresponds to the j^{th} one of the modes (traveling in the positive z direction) for which A is proportional to $\cos(\pi \ell y/b)$.

In much the same way it may be shown that the series (1.2-8) assumed for B is a solution of equation (1.2-6) (with A replaced by B) provided the coefficients $\beta_{mn}(z)$ satisfy the set of equations

$$-\Gamma_p^2 \beta_{p\ell}(z) + \sum_{m=0}^{\infty} [Q_{pm} \beta_m''(z) - U_{pm} \beta_m(z)] = 0 \quad (1.2-14)$$

for $p = 0, 1, 2, \dots$ and $\ell = 1, 2, 3, \dots$. Here

$$Q_{pm} = (\epsilon_p/a) \int_0^a (\rho_1^2/\rho^2) \cos(\pi p x/a) \cos(\pi m x/a) dx \quad (1.2-15)$$

$$U_{pm} = \pi m \epsilon_p a^{-2} \int_0^a \cos(\pi p x/a) \sin(\pi m x/a) dx / \rho \quad (1.2-16)$$

where $\epsilon_0 = 1$ and $\epsilon_p = 2$ for $p > 0$. These integrals are discussed in Appendix I.

The problem of determining the reflection from a bend in a wave guide involves considerable manipulation of equations (1.2-9) and (1.2-14). The introduction of matrix notation in the manner suggested by the work of Section 1.1 for straight guides simplifies this work. Although $\alpha_{mn}(z)$ and $\beta_{mn}(z)$ are no longer the simple exponential functions given by (1.1-9), it turns out that the column matrices $\alpha(z)$ and $\beta(z)$ are still given by (for a wave traveling in the positive z direction) by the matrix expression (1.1-10). As mentioned earlier, Γ_α and Γ_β are no longer simple diagonal matrices.

We now turn to the task of expressing (1.2-9) and (1.2-14) in matrix form.

1.3 Propagation Constant Matrix for Curved Rectangular Guide

From this point onward in our investigation of propagation in the rectangular guide we shall assume A to be proportional to $\cos(\pi\ell y/b)$. Thus instead of the general expression (1.2-7) for A we shall deal with the more restricted form

$$A = \cos(\pi\ell y/b) \sum_{m=1}^{\infty} \alpha_{m\ell}(z) \sin(\pi mx/a) \quad (1.3-1)$$

where ℓ has one of the values $0, 1, 2, 3, \dots$. Since the most general disturbance may be obtained by the superposition of disturbances of the form (1.3-1) no real generality will be lost.

The introduction of (1.3-1) is suggested by the fact that the set $\alpha_{1\ell}(z)$, $\alpha_{2\ell}(z)$, $\dots$ may be determined from (1.2-9) (at least to within arbitrary constants of integration) without considering the other $\alpha_{mn}(z)$'s, $n \neq \ell$. The introduction of (1.3-1) is also suggested by physical reasons. The plane of the bend is the z, x plane and there is nothing in the system tending to change the field distribution in the y direction.

Equation (1.2-13) is a special case of (1.3-1). Furthermore the most general form of (1.3-1) (corresponding to a wave progressing in the positive z direction) may be obtained by multiplying (1.2-13) by an arbitrary constant c_j and summing on j .

In order to write the set of differential equations (1.2-9) for the $\alpha_{m\ell}(z)$'s in matrix form we introduce the infinite matrices

$$\Gamma_0 = \begin{bmatrix} \Gamma_{1\ell} & 0 & 0 & \cdot \\ 0 & \Gamma_{2\ell} & 0 & \cdot \\ 0 & 0 & \Gamma_{3\ell} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}, \quad \alpha(z) = \begin{bmatrix} \alpha_{1\ell}(z) \\ \alpha_{2\ell}(z) \\ \alpha_{3\ell}(z) \\ \cdot \end{bmatrix} \quad (1.3-2)$$

$$P = \begin{bmatrix} P_{11} & P_{12} & \cdot \\ P_{21} & P_{22} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}, \quad S = \begin{bmatrix} S_{11} & S_{12} & \cdot \\ S_{21} & S_{22} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

where the elements of Γ_0 are obtained by setting $n = \ell$ in equation (1.1-5) for Γ_{mn} , and the elements of P and Q are given by the integrals (1.2-10) and (1.2-11). The rules of matrix multiplication then show that (1.2-9) is the p^{th} element of the matrix equation

$$P\alpha''(z) - (\Gamma_0^2 + S)\alpha(z) = 0 \quad (1.3-3)$$

Premultiplying by P^{-1} converts this equation into

$$\alpha''(z) - \Gamma_\alpha^2 \alpha(z) = 0 \quad (1.3-4)$$

where

$$\Gamma_{\alpha}^2 = P^{-1} (\Gamma_0^2 + S) \quad (1.3-5)$$

It may be verified by direct differentiation of the series (1.1-11) defining $\exp(-z\Gamma)$ that a solution of (1.3-4) is

$$\alpha(z) = e^{-z\Gamma_{\alpha}} g^+ \quad (1.3-6)$$

where, as in (1.1-10) for the straight guide, g^+ is a column of constants (of integration). However, now Γ_{α} is to be obtained by taking the square root of the right hand side of (1.3-5), a process which is not easy since it usually requires one to obtain the characteristic roots and modal columns of Γ_{α}^2 (see equation (1.3-10)).

As far as (1.3-6) being a solution of the differential equation is concerned, Γ_{α} may be any matrix whose square is given by (1.3-5). We shall restrict it as follows: When $\xi = a/\rho_1$ becomes small, as in the case of a gentle bend, it is seen from (1.2-10) and (1.2-11) that P approaches the unit matrix and S approaches zero. Hence, Γ_{α}^2 approaches the diagonal matrix Γ_0^2 . Γ_{α} is chosen so that it approaches Γ_0 , that is, all of the elements in the principal diagonal are either positive real or positive imaginary. This makes $\exp(-z\Gamma_{\alpha})$ approach the diagonal matrix $\exp(-z\Gamma_0)$. With this choice of Γ_{α} the expression (1.3-6) for $\alpha(z)$ corresponds to a wave traveling in the positive z direction.

The various modes of propagation in the bend may be obtained from Γ_{α}^2 by expressing, in matrix notation, the steps leading to (1.2-13) (which gives A for the j^{th} mode). We assume $\alpha(z)$ to be the column matrix obtained by multiplying the column matrix g of constants by the scalar quantity $e^{\gamma z}$. Setting this in (1.3-4) gives

$$(\gamma^2 I - \Gamma_{\alpha}^2)g = 0 \quad (1.3-7)$$

where I is the unit matrix. In order that (1.3-7) may have a solution, the determinant of the coefficient of g must vanish. This leads to the characteristic equation* for γ^2 :

$$|\gamma^2 I - \Gamma_{\alpha}^2| = 0 \quad (1.3-8)$$

The vertical bars denote the determinant of the inclosed matrix. The roots $\gamma_1^2, \gamma_2^2, \dots$ are therefore the latent (or characteristic) roots of Γ_{α}^2 . If we let k_j denote** the column g obtained when $\gamma = \gamma_j$ in (1.3-7) then

* See Section 3.6 of Reference⁹.

** We choose this notation in order to adhere as closely as possible to that of Reference⁹. Incidentally, the column k_j is proportional to the j^{th} column of κ^{-1} where κ is the modal row matrix introduced in Section 5.1.

$$(\gamma_j^2 I - \Gamma_\alpha^2) k_j = 0 \quad (1.3-9)$$

and the elements $k_{1j}, k_{2j}, \dots$ of the modal column k_j are the ones appearing as coefficients in (1.2-13).

Equation (1.3-9) and the methods of matrix analysis lead to

$$\Gamma_\alpha^2 = k[\gamma^2]_d k^{-1}, \quad \Gamma_\alpha = k[\gamma]_d k^{-1} \quad (1.3-10)$$

where k is the square matrix whose j^{th} column is k_j and $[\gamma^2]_d, [\gamma]_d$ are diagonal matrices having γ_j^2, γ_j as the j^{th} elements in their principal diagonals. The representation (1.3-10) certainly holds for the rectangular guide since in this case no repeated roots occur.

In analogy with the expression (1.3-1) for A we shall henceforth deal with B in the form

$$B = \sin(\pi \ell y/b) \sum_{m=0}^{\infty} \beta_{m\ell}(z) \cos(\pi m x/a) \quad (1.3-11)$$

where ℓ has one of the values $1, 2, 3, \dots$. In much the same way as before it may be shown that for a wave traveling along the bend in the positive direction the $\beta_{m\ell}(z)$'s in (1.3-11) are given by

$$\beta(z) = e^{-z\Gamma_\beta} d^+ \quad (1.3-12)$$

where d^+ is a column of arbitrary constants and

$$\Gamma_\beta^2 = Q^{-1}(\Gamma_0^2 + U) \quad (1.3-13)$$

In (1.3-12) and (1.3-13)

$$\Gamma_0 = \begin{bmatrix} \Gamma_{0\ell} & 0 & 0 & \cdot \\ 0 & \Gamma_{1\ell} & 0 & \cdot \\ 0 & 0 & \Gamma_{2\ell} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad \beta(z) = \begin{bmatrix} \beta_{0\ell}(z) \\ \beta_{1\ell}(z) \\ \beta_{2\ell}(z) \\ \cdot \end{bmatrix} \quad (1.3-14)$$

$$Q = \begin{bmatrix} Q_{00} & Q_{01} & \cdot \\ Q_{10} & Q_{11} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad U = \begin{bmatrix} 0 & U_{01} & U_{02} & \cdot \\ 0 & U_{11} & U_{12} & \cdot \\ 0 & U_{21} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

where the elements of Γ_0, Q and U are given by equations (1.1-5), (1.2-15) and (1.2-16), respectively.

1.4 Continuity Conditions at Junction of Straight and Curved Rectangular Guides

Electromagnetic theory requires that E_x, E_y, H_x and H_y be continuous in crossing a plane $z = \text{constant}$ which marks the junction of a straight and a

curved wave guide (of the same cross-section). Comparison of the first equation in (1.1-7) with the first equation in (1.2-5) shows that E_x is continuous if (1) A is continuous and (2) if $\partial B/\partial z$ in the straight portion is equal (the equality being taken at the junction) to $(\rho_1/\rho) \partial B/\partial z$ in the curved portion. Examination of the expressions for the remaining field components shows that all the continuity conditions are satisfied if, at the junction,

$$[A \text{ in straight portion}] = [A \text{ in bend}]$$

$$\left[\frac{\partial A}{\partial z} \quad " \quad " \quad " \right] = \left[\frac{\rho_1}{\rho} \frac{\partial A}{\partial z} \text{ in bend} \right] \quad (1.4-1)$$

and likewise for B .

Let A in the bend be given by (1.3-1) and let $\alpha(z)$ denote the column matrix of coefficients shown in (1.3-2). A in the straight portion may be represented in the same way except that $\alpha(z)$ has a simpler form as explained in Section 1.1. When these expressions for A are inserted in (1.4-1), both sides multiplied by $(2/a)\sin(\pi px/a)$ after cancelling out the $\cos(\pi ly/b)$, and the results integrated with respect to x from 0 to a we obtain relations which may be expressed as the matrix equations

$$[\alpha(z) \text{ in straight portion}] = [\alpha(z) \text{ in bend}]$$

$$\left[\frac{d\alpha(z)}{dz} \quad " \quad " \quad " \right] = \left[V \frac{d\alpha(z)}{dz} \text{ in bend} \right] \quad (1.4-2)$$

where V is the square matrix whose p^{th} row and m^{th} column ($p, m = 1, 2, 3, \dots$) is

$$V_{pm} = (2\rho_1/a) \int_0^a \sin(\pi px/a) \sin(\pi mx/a) dx / \rho, \quad (1.4-3)$$

ρ being equal to $\rho_1 + x - a/2$.

By using expression (1.3-11) for B in the continuity conditions, it may be shown in much the same way that the column matrix $\beta(z)$ given by (1.3-14) must satisfy the relations

$$[\beta(z) \text{ in straight portion}] = [\beta(z) \text{ in bend}]$$

$$\left[\frac{d\beta(z)}{dz} \quad " \quad " \quad " \right] = \left[W \frac{d\beta(z)}{dz} \text{ in bend} \right] \quad (1.4-4)$$

where W is the infinite square matrix

$$W = \begin{bmatrix} W_{00} & W_{01} & \cdot \\ W_{10} & W_{11} & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \quad (1.4-5)$$

whose elements are

$$W_{pm} = (\epsilon_p \rho_1/a) \int_0^a \cos(\pi p x/a) \cos(\pi m x/a) dx/\rho \quad (1.4-6)$$

$$\epsilon_0 = 1, \quad \epsilon_p = 2 \quad \text{for } p > 0$$

Both V_{pm} and W_{pm} are discussed in Appendix I.

PART II

THEORY FOR A GENERAL WAVE GUIDE

2.1 Matrix Propagation Constant for a Curved Wave Guide of Arbitrary Cross-Section

In Section 1.3 it has been shown that for a curved rectangular wave guide there exists a square matrix Γ_α (or Γ_β) which plays the same role in the propagation of a wave consisting of many modes as does the propagation constant in a simple transmission line. There Γ_α was obtained from a special form of the wave equation which is suited to bends in rectangular guides. Here we adopt a different approach with the idea of showing that a matrix propagation constant Γ exists under more general conditions.

The general theory of wave propagation in tubes shows that a wave traveling in the positive z direction may often be represented as

$$\Phi = \sum_{j=1}^{\infty} c_j e^{-z\gamma_j} \varphi_j(x, y) \quad (2.1-1)$$

where Φ is some quantity associated with the field and is analogous to the functions A and B of Part I. In (2.1-1) x and y are transverse coordinates and z a longitudinal coordinate. γ_j is the propagation constant for the j^{th} mode and $\varphi_j(x, y)$ the corresponding eigenfunction. For a circular bend in a rectangular wave guide $\varphi_j(x, y)$ is a combination of trigonometric and Bessel functions and γ_j is proportional to the order of the Bessel functions.

We assume that we may find a set of functions $\theta_m(x, y)$, $m = 1, 2, 3, \dots$ such that every $\varphi_j(x, y)$ may be represented as

$$\varphi_j(x, y) = \sum_{m=1}^{\infty} k_{mj} \theta_m(x, y) \quad (2.1-2)$$

The usefulness of this procedure depends upon our ability to pick a system of $\theta_m(x, y)$'s which is appreciably simpler than the system of $\varphi_j(x, y)$'s. In the work of Part I $\theta_m(x, y)$ was taken to be the eigenfunction of the typical mode of propagation in a straight guide, i.e. the product of a sine and a cosine.

We assume further in (2.1-2) that the square matrix k^{-1} exists where k_{mj} is

the element in the m^{th} row and j^{th} column of k ; i.e. if a root γ_j is repeated, say, s times there are s linearly independent columns (k_j 's) corresponding to γ_j . Substitution of (2.1-2) in (2.1-1) gives

$$\begin{aligned}\Phi &= \sum_{m=1}^{\infty} \theta_m(x, y) \sum_{j=1}^{\infty} k_{mj} c_j e^{-z\gamma_j} \\ &= \sum_{m=1}^{\infty} \mu_m(z) \theta_m(x, y)\end{aligned}\quad (2.1-3)$$

where

$$\mu_m(z) = \sum_{j=1}^{\infty} k_{mj} c_j e^{-z\gamma_j} \quad (2.1-4)$$

Since $\theta_m(x, y)$ is analogous to the product of the trigonometrical terms in (1.3-1) or (1.3-11) these equations show that $\mu_m(z)$ plays the same role as $\alpha_m \ell(z)$ or $\beta_m \ell(z)$. Therefore, in accordance with the discussion given at the beginning of this section, we wish to show that the column matrix $\mu(z)$ (which is similar to $\alpha(z)$ or $\beta(z)$) whose m^{th} element is $\mu_m(z)$ may be expressed as

$$\mu(z) = e^{-z\Gamma} f^+ \quad (2.1-5)$$

In this equation Γ is a square matrix to be determined and f^+ is a column matrix of constants similar to g^+ or d^+ .

The rules of matrix multiplication and equation (2.1-4) show that

$$\mu(z) = k[e^{-z\gamma}]_d c \quad (2.1-6)$$

in which $[\exp(-z\gamma)]_d$ is a diagonal matrix having $\exp(-z\gamma_j)$ as the j^{th} element in its principal diagonal and c is the column matrix formed from the c_j 's. We introduce the column f^+ by defining it as $\mu(0)$ whence

$$f^+ = kc, \quad c = k^{-1}f^+ \quad (2.1-7)$$

Incidentally, from (2.1-3), the value of Φ at $z = 0$ is

$$\Phi_{z=0} = \sum_{m=1}^{\infty} f_m^+ \theta_m(x, y) \quad (2.1-8)$$

where f_m^+ is the m^{th} element in f^+ .

From (2.1-6) and (2.1-7)

$$\mu(z) = k[e^{-z\gamma}]_d k^{-1} f^+ \quad (2.1-9)$$

In this equation $k[\exp(-z\gamma)]_d k^{-1}$ is a square matrix which may be expressed as

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{(-z)^n}{n!} k[\gamma]_d^n k^{-1} &= \sum_{n=0}^{\infty} \frac{(-z)^n}{n!} (k[\gamma]_d k^{-1})^n \\ &= e^{-z\Gamma}\end{aligned}\quad (2.1-10)$$

Here $[\gamma]_d$ represents the diagonal matrix having γ_i the j^{th} term in its principal diagonal and

$$\Gamma = k[\gamma]_d k^{-1} \quad (2.1-11)$$

Therefore we have shown that $\mu(z)$ is of the form (2.1-5) which is what we set out to do.

It is rather difficult to compute Γ from (2.1-11) using only the above definitions of k and γ_i for one must first obtain the functions $\varphi_i(x, y)$. In dealing with the rectangular guide it is easier to use equations (1.3-5) and (1.3-13) to determine Γ .

2.2 Reflection at a Single Junction

Let a straight wave guide extending from $z = -\infty$ to $z = 0$ be joined to a curved guide of the same cross-section which extends from $z = 0$ to $z = \infty$. Let an incident wave

$$\Phi_i = \sum_{m=1}^{\infty} h_m e^{-z\delta_m} \theta_m(x, y) \quad (2.2-1)$$

come in from the left along the straight guide. The h_m 's are given constants, the δ_m 's are the modal propagation constants for straight guides (for rectangular guides they are the Γ_m 's given by (1.1-5)), and $\theta_m(x, y)$ is the m^{th} eigenfunction for the straight guide (the product of a sine and a cosine for a rectangular guide).

What are the reflected and transmitted waves set up by (2.2-1)? The reflected wave is of the form

$$\Phi_r = \sum_{m=1}^{\infty} f_m^- e^{z\delta_m} \theta_m(x, y) \quad (2.2-2)$$

where the f_m^- 's are to be determined.

If we assume the representation

$$\Phi = \sum_{m=1}^{\infty} \mu_m(z) \theta_m(x, y) \quad (2.2-3)$$

to hold for all real values of z then, since $\Phi = \Phi_i + \Phi_r$ for $z < 0$, equations (2.2-1) and (2.2-2) show that

$$\mu_m(z) = h_m e^{-z\delta_m} + f_m^- e^{z\delta_m}, \quad m = 1, 2, 3, \dots; z < 0 \quad (2.2-4)$$

Introducing the column matrices $\mu(z)$, h , f^- and the diagonal matrices $\exp(\pm z\Gamma_0)$ where Γ_0 is a diagonal matrix having δ_m as the m^{th} term in its principal diagonal enables us to write (2.2-4) as

$$\mu(z) = e^{-z\Gamma_0} h + e^{z\Gamma_0} f^-, \quad z < 0 \quad (2.2-5)$$

Equation (2.2-5) is more general than the expression (2.1-5) for $\mu(z)$ in that it contains waves going in both directions, but is more special in that Γ_0 is a diagonal matrix.

In the curved guide we take $\mu(z)$ to be given by (2.1-5), thus

$$\mu(z) = e^{-z\Gamma} f^+, \quad z > 0 \quad (2.2-6)$$

where Γ is a square matrix whose elements are assumed to be known and f^+ is a column matrix whose elements are to be determined along with those of f^- .

The conditions that the transverse components of the electric and magnetic intensities be continuous at the junction of the two guides are assumed to lead to the requirements

$$[\mu(z) \text{ in straight portion}] = [\mu(z) \text{ in curved portion}]$$

$$\left[\frac{d}{dz} \mu(z) \text{ in straight portion} \right] = \left[V \frac{d}{dz} \mu(z) \text{ in curved portion} \right] \quad (2.2-7)$$

where the quantities within the brackets are evaluated at the junction and V is a square matrix whose elements are constants. When the curvature of the curved portion becomes small V approaches the unit matrix. For the problem at hand (2.2-7) may be written as

$$[\mu(z)]_{z=-0} = [\mu(z)]_{z=+0} \quad (2.2-8)$$

$$\left[\frac{d}{dz} \mu(z) \right]_{z=-0} = V \left[\frac{d}{dz} \mu(z) \right]_{z=+0} \quad (2.2-9)$$

in which the subscripts $z = -0$, $z = +0$ refer to the straight and curved portions, respectively, of the guide at $z = 0$.

The requirements (2.2-7) have been established for the rectangular guide in Section 1.4. Their form is also suggested by the conditions that the voltage and current be continuous at the junction of two transmission lines. Thus if we let $\mu(z)$ play the role of the voltage, the current in the first line is $-Z_1^{-1} d\mu(z)/dz$ and the current in the second is $-Z_2^{-1} d\mu(z)/dz$ where Z_1 and Z_2 denote the distributed series impedances of the two lines. It is seen that this leads to scalar equations which look like (2.2-7), but now V denotes the scalar Z_1/Z_2 instead of a square matrix.

Setting the expressions (2.2-5) and (2.2-6) for $\mu(z)$ in the conditions (2.2-8) and (2.2-9) gives two equations which may be solved simultaneously to obtain f^- and f^+ in terms of h , Γ_0 , Γ and V :

$$h + f^- = f^+ \quad (2.2-10)$$

$$-\Gamma_0 h + \Gamma_0 f^- = -V\Gamma f^+$$

$$f^- = (\Gamma_0 + V\Gamma)^{-1}(\Gamma_0 - V\Gamma)h \quad (2.2-11)$$

$$f^+ = (\Gamma_0 + V\Gamma)^{-1}\Gamma_0 h \quad (2.2-12)$$

Since f^- and f^+ specify the reflected and transmitted waves, respectively, they give the answer which we are seeking.

If the curved guide should extend from $z = -\infty$ to $z = 0$ and the straight guide from $z = 0$ to $z = \infty$ the response to an incident wave $e^{-z\Gamma} h$ coming in along the curved guide would be

$$\begin{aligned}\mu(z) &= e^{-z\Gamma} h + e^{z\Gamma} f^-, & z < 0 \\ \mu(z) &= e^{-z\Gamma_0} f^+, & z > 0\end{aligned}\quad (2.2-13)$$

A procedure similar to that used above shows that

$$\begin{aligned}f^- &= -(\Gamma_0 + V\Gamma)^{-1} (\Gamma_0 - V\Gamma)h \\ f^+ &= (\Gamma_0 + V\Gamma)^{-1} 2V\Gamma h\end{aligned}\quad (2.2-14)$$

where, instead of condition (2.2-9), we have used

$$V \left[\frac{d}{dz} \mu(z) \right]_{z=-0} = \left[\frac{d}{dz} \mu(z) \right]_{z=+0} \quad (2.2-15)$$

2.3 Reflection Due to a Bend

Let the guide be straight for $-\infty < z < -c$ and for $c < z < \infty$, and let these two portions be connected by a curved portion in which the longitudinal coordinate z runs from $-c$ to $+c$. As in Section 2.2 we take the matrix propagation constants for the straight and curved portions to be the square matrices Γ_0 and Γ , respectively, and assume an incident wave, specified by the column matrix h , to come in from $z = -\infty$.

The column matrix $\mu(z)$ whose m^{th} element appears as the coefficient of $\theta_m(x, y)$ in the representation (2.2-3) for Φ is now given by

$$\begin{aligned}\mu(z) &= e^{-z\Gamma_0} h + e^{z\Gamma_0} f^-, & z < -c \\ \mu(z) &= (\cosh z\Gamma)p + (\sinh z\Gamma)q, & -c < z < c \\ \mu(z) &= e^{-z\Gamma_0} f^+, & c < z\end{aligned}\quad (2.3-1)$$

In these expressions f^-, f^+, p, q are column matrices which may be determined as functions of the known matrices Γ_0, Γ, V and h by substituting (2.3-1) in the conditions (2.2-7) which must hold at the junctions $z = -c$ and $z = c$.

By straightforward algebra similar to that used for the analogous problem in transmission line theory we obtain

$$\begin{aligned}e^{-c\Gamma_0} f^- + e^{-c\Gamma_0} f^+ &= [-I + 2(V\Gamma \tanh c\Gamma + \Gamma_0)^{-1} \Gamma_0] e^{c\Gamma_0} h \\ e^{-c\Gamma_0} f^- - e^{-c\Gamma_0} f^+ &= [-I + 2(V\Gamma \coth c\Gamma + \Gamma_0)^{-1} \Gamma_0] e^{c\Gamma_0} h\end{aligned}\quad (2.3-2)$$

In these equations the infinite square matrix $\tanh c\Gamma$ is defined as $(\sinh c\Gamma)/(\cosh c\Gamma)^{-1}$ and $\coth c\Gamma$ as its reciprocal. $\sinh c\Gamma$ and $\cosh c\Gamma$ may be defined as power series in $c\Gamma$ and may be expressed as combinations of $\exp(c\Gamma)$ and $\exp(-c\Gamma)$.

An expression for the column matrix f^- may be obtained by adding the equations in (2.3-2). Before doing this it is convenient to introduce the two column matrices x and y defined by

$$\begin{aligned}(Vc\Gamma \tanh c\Gamma + c\Gamma_0) x &= c\Gamma_0 e^{c\Gamma_0} h \\ (Vc\Gamma \coth c\Gamma + c\Gamma_0) y &= c\Gamma_0 e^{c\Gamma_0} h\end{aligned}\tag{2.3-3}$$

where the scalar length c has been introduced to make the various terms dimensionless. Each equation in (2.3-3) represents an infinite set of simultaneous linear equations to be solved for the elements of x or y .

Once x and y are known the reflected wave is given by

$$f^- = e^{c\Gamma_0} (x + y) - e^{2c\Gamma_0} h\tag{2.3-4}$$

and the transmitted wave by

$$f^+ = e^{c\Gamma_0} (x - y)\tag{2.3-5}$$

PART III

GENTLE BENDS—GENERAL THEORY

3.1 Limiting Forms Assumed for Γ and V

It will be shown in Part IV that for gentle circular bends in rectangular wave guides the matrix propagation constant Γ is such that

$$\Gamma^2 = \Gamma_0^2 + F\tag{3.1-1}$$

where Γ_0^2 is the square of the matrix propagation constant for the straight guide. Γ_0^2 is a diagonal matrix having δ_m^2 (which is one of the Γ_{mn}^2 's, depending on the set of modes under consideration, given by (1.1-5)) for the m^{th} element in its principal diagonal. F is a square matrix of infinite order in which the elements F_{ii} in the principal diagonal are of order ξ^2 and the remaining elements F_{ij} , $i \neq j$ are of order ξ . Here $\xi = a/\rho_1$ is the ratio of the guide width to the radius of curvature of the bend. As the bend becomes more and more gentle, $\xi \rightarrow 0$.

The asymptotic expressions given in Appendix I show that, for gentle bends in rectangular guides, the square matrix V which appears in the junction conditions (2.2-7) approaches a unit matrix as $\xi \rightarrow 0$. In particular $V_{ii} = 1 + v_{ii}$ where v_{ii} is of order ξ^2 , and V_{ij} , the element in the i^{th} row and j^{th} column, is of order ξ when $i \neq j$.

Throughout the remainder of Part III we shall assume that Γ^2 , F and V behave as mentioned above. In addition we assume that there is no degeneracy, i.e. all of the δ_m 's are unequal to each other and to zero.

3.2 Propagation in a Gentle Bend

Here we assume that the elements of Γ_0^2 and F in the expression (3.1-1) for Γ^2 are known. We wish to find the modal propagation constant γ_j and the corresponding eigenfunction $\varphi_j(x, y)$ for the j^{th} mode.

After squaring both sides of the collineatory transformation (2.1-11) connecting Γ and the diagonal matrix $[\gamma]_d$ we obtain a relation which may be written as $k[\gamma^2]_d - \Gamma^2 k = 0$. The left hand side is a square matrix having $(\gamma_j^2 I - \Gamma^2)k_j$ as its j^{th} column. Here I is the unit matrix and k_j is a column matrix having $k_{1j}, k_{2j}, \dots$ as its elements (k_j is the j^{th} column of k). Thus we have a system of simultaneous linear equations in which the coefficients are furnished by the square matrix $\gamma_j^2 I - \Gamma^2$ and in which the unknowns are $k_{1j}, k_{2j}, \dots$. Accordingly, γ_j^2 is the j^{th} latent root of Γ^2 and k_j is its corresponding modal column just as for the rectangular guide in Section 1.3.

In order to apply equations (A2-16) of Appendix II we set $\lambda_j = \gamma_j^2$ and $u = \Gamma^2$ so that, from (3.1-1),

$$u_{jj} = \delta_j^2 + F_{jj}, \quad u_{ij} = F_{ij}, \quad i \neq j \quad (3.2-1)$$

Therefore

$$\gamma_j^2 = \delta_j^2 + F_{jj} + \sum_{s=1}^{\infty} F_{js} F_{sj} / (\delta_j^2 - \delta_s^2) \quad (3.2-2)$$

$$k_{jj} = 1, \quad k_{sj} = F_{sj} / (\delta_j^2 - \delta_s^2), \quad s \neq j \quad (3.2-3)$$

where we have neglected terms of order ξ^3 in (3.2-2) and of order ξ^2 in k_{sj} , $s \neq j$. The prime on the summation indicates that the term $s = j$ is to be omitted.

When $k_{1j}, k_{2j}, \dots$ are known the eigenfunction $\varphi_j(x, y)$ may be written as a series in $\theta_m(x, y)$ by means of equation (2.1-2).

In Section 3.3 we shall need the form assumed by the square matrix $\Gamma \tanh c\Gamma$ in a gentle bend. This matrix is used in computing the reflection from such a bend, as might be inferred from equation (2.3-3). The formula to be used is (A2-18) with $u = \Gamma^2$, $\lambda_j = \gamma_j^2$ and with the elements of the square matrix k given by (3.2-3). In the diagonal matrix of (A2-18) we set

$$f(\lambda_j) = \lambda_j^{1/2} \tanh c\lambda_j^{1/2} = \gamma_j \tanh c\gamma_j = \gamma_j t_j \quad (3.2-4)$$

$$t_j = \tanh c\gamma_j$$

and for the elements of k^{-1} we use (A2-19) together with the line above it. When the three matrices on the right of (A2-18) are multiplied out the element $(\Gamma \tanh c\Gamma)_{ij}$ in the i^{th} row and j^{th} column of $\Gamma \tanh c\Gamma$ is found to be

$$(\Gamma \tanh c\Gamma)_{ij} = (\gamma_{itj} - \gamma_{iti})k_{iji}, \quad i \neq j \quad (3.2-5)$$

$$(\Gamma \tanh c\Gamma)_{ii} = \gamma_{iti} + \sum'_m (\gamma_{it_i} - \gamma_{mtm})k_{im}k_{mi} \quad (3.2-6)$$

where in (3.2-5) and (3.2-6) terms of $O(\xi^2)$ ("order of") and $O(\xi^3)$, respectively, have been neglected. This is in line with the fact that the terms in the principal diagonals of our matrices must be accurate to within $O(\xi^2)$ while the remaining terms need be accurate only to within terms of $O(\xi)$. The summation with respect to m runs from $m = 1$ to ∞ with $m = i$ omitted.

3.3 Reflection from a Gentle Bend

When the bend is gentle so that V and Γ behave according to the description given in Section 3.1, the matrix expressions for the reflection coefficients given in Section 2.3 may be evaluated. The results stated in Appendix II for "almost diagonal" matrices furnish the principal tools for this work.

It is assumed that the incident wave coming in along the straight guide from the left is $\Phi_i = \exp(-z\delta_p) \theta_p(x, y)$ and hence contains only the p^{th} mode. Comparing this with the general expression (2.2-1) for Φ_i shows that $h_p = 1$, $h_m = 0$, $m \neq p$, and all the elements of the column matrix h are zero except the p^{th} which is unity.

We start by writing the first of equations (2.3-3) as

$$(\Gamma \tanh c\Gamma + V^{-1}\Gamma_0)x = V^{-1}\Gamma_0 e^{c\Gamma_0} h \quad (3.3-1)$$

Since V approaches a unit matrix as $\xi \rightarrow 0$, the element $(V^{-1})_{ij}$ in the i^{th} row and j^{th} column of V^{-1} is

$$\begin{aligned} (V^{-1})_{ij} &= -V_{iji}, \quad i \neq j \\ (V^{-1})_{ii} &= 1 - v_{ii} + \sum'_m V_{im} V_{mi} \end{aligned} \quad (3.3-2)$$

where $V_{ii} = 1 + v_{ii}$, $i, j = 1, 2, 3, \dots$ and the summation with respect to m runs from 1 to ∞ with the term for $m = i$ omitted (as indicated by the prime on Σ). In omitting this term we are neglecting v_{ii}^2 because it is of order ξ^4 . These results follow from equation (A2-2) of Appendix II. As usual, the elements in the principal diagonal are accurate to within $O(\xi^2)$ and the remaining elements to within $O(\xi)$.

It follows that $V^{-1}\Gamma_0 e^{c\Gamma_0} h = \eta$ is a column matrix whose i^{th} element is

$$\begin{aligned} \eta_i &= -V_{ip} \delta_p e^{c\delta_p}, \quad i \neq p \\ \eta_p &= (1 - v_{pp} + \sum'_m V_{pm} V_{mp}) \delta_p e^{c\delta_p}, \quad i = p \end{aligned} \quad (3.3-3)$$

Likewise, the element in the i^{th} row and j^{th} column of the square matrix $V^{-1}\Gamma_0$ is $-V_{ij}\delta_j$ when $i \neq j$ and is

$$(1 - v_{ii} + \sum'_m V_{im} V_{mi}) \delta_i \quad (3.3-4)$$

when $i = j$.

By combining the approximate expressions for the elements of $\Gamma \tanh c\Gamma$ (given by (3.2-5) and (3.2-6)) and $V^{-1}\Gamma_0$ we find that if u_{ij} denotes the i^{th} row and j^{th} column of $\Gamma \tanh c\Gamma + V^{-1}\Gamma_0$ then

$$\begin{aligned} u_{ij} &= D_{ij} - V_{ij}\delta_j, \quad i \neq j \\ u_{ii} &= \gamma_i t_i + \sum'_m D_{mi} k_{im} + \delta_i(1 - v_{ii} + \sum'_m C_{mi}) \\ &= \sigma_i - \delta_i v_{ii} + \sum'_m (D_{mi} k_{im} + \delta_i C_{mi}) \end{aligned} \quad (3.3-5)$$

In these equations we have set

$$\begin{aligned} \sigma_i &= \delta_i + \gamma_i t_i, \quad C_{mi} = V_{im} V_{mi} \\ D_{mi} &= (\gamma_i t_i - \gamma_m t_m) k_{mi} = (\gamma_i t_i - \gamma_m t_m) F_{mi} / (\delta_i^2 - \delta_m^2) \end{aligned} \quad (3.3-6)$$

where γ_i and k_{mi} are given by (3.2-2) and (3.2-3).

We are now in a position to identify the matrix equation (3.3-1) for x with the set of equations (A2-20). The quantity η_i which appears on the right hand side of the i^{th} equation in (A2-20) is given by (3.3-3). The coefficients which appear on the left hand side are the u 's defined by (3.3-5). Therefore, from (A2-21), when $i \neq p$,

$$\begin{aligned} x_i &= \frac{\delta_p e^{c\delta_p}}{\sigma_i \sigma_p} [-V_{ip}(\sigma_p - \delta_p) - D_{ip}] \\ &= -\frac{e^{c\delta_p}}{\delta_i(1+t_i)(1+t_p)} [V_{ip}\delta_p t_p + F_{ip}(\delta_i t_i - \delta_p t_p)(\delta_i^2 - \delta_p^2)^{-1}] \end{aligned} \quad (3.3-7)$$

where we have neglected higher order terms and in so doing have replaced γ_j by the simpler δ_j .

When $i = p$ (A2-21) yields

$$x_p = u_{pp}^{-1}[\eta_p + \sum'_m (u_{pm} u_{mp} \eta_p - u_{pm} u_{pp} \eta_m) / (u_{mm} u_{pp})] \quad (3.3-8)$$

In order to combine the second order terms in $1/u_{pp}$ with those in the rest of the expression for x_p we assume that σ_p is the major portion of u_{pp} . Then, approximately,

$$1/u_{pp} = \sigma_p^{-1} [1 + \delta_p v_{pp} / \sigma_p - \sum'_m (D_{mp} k_{pm} + \delta_p C_{mp}) / \sigma_p] \quad (3.3-9)$$

This assumption, which is equivalent to assuming that $\tanh c\gamma_p$ differs appreciably from -1 , does not appear to restrict our results since $\tanh c\gamma_p$ is either purely imaginary or else real and positive.

Substituting the appropriate values in (3.3-8), neglecting higher order terms, and using the definition (3.3-6) for C_{mp} leads to

$$x_p = \sigma_p^{-1} \delta_p e^{c\delta_p} [1 - (1 - \delta_p/\sigma_p) v_{pp} + \sum'_m (\sigma_p \sigma_m)^{-1} \{ C_{mp}(\sigma_p - \delta_p)(\sigma_m - \delta_m) + D_{mp}(D_{pm} - \sigma_m k_{pm}) + (\sigma_p - \delta_p) D_{pm} V_{mp} - \delta_m D_{mp} V_{pm} \}] \quad (3.3-10)$$

A reduction similar to that used in going from the first to the second line of (3.3-7) gives our final expression for x_p

$$x_p = \frac{\delta_p e^{c\delta_p}}{\delta_p + \gamma_p t_p} - \frac{t_p e^{c\delta_p}}{(1 + t_p)^2} \left[v_{pp} - \sum'_m V_{pm} V_{mp} t_m / (1 + t_m) + \sum'_m \frac{\delta_m t_m - \delta_p t_p}{\delta_m \delta_p (1 + t_m) (\delta_m^2 - \delta_p^2)} \{ V_{pm} \delta_m F_{mp} / t_p - V_{mp} \delta_p F_{pm} + F_{pm} F_{mp} (\delta_p + \delta_m / t_p) (\delta_m^2 - \delta_p^2)^{-1} \} \right] \quad (3.3-11)$$

The above expressions for x_i and x_p have been derived from the first of equations (2.3-3). The second of equations (2.3-3) determines the column matrix y in the same way that the first equation determines x except that $\coth c\Gamma$ now replaces $\tanh c\Gamma$. Therefore, we may obtain expressions for the elements of y by replacing the t 's (where $t_i = \tanh c\gamma_i$) by their reciprocals in the expressions for the corresponding x 's (i.e. in (3.3-7) and (3.3-11). The values obtained in this way lead to, when $i \neq p$,

$$\begin{aligned} f_i^- &= e^{c\delta_i} (x_i + y_i) \\ &= \delta_i^{-1} e^{c(\delta_i + \delta_p - \gamma_i - \gamma_p)} \sinh c(\gamma_i + \gamma_p) [-V_{ip} \delta_p - F_{ip} / (\delta_i + \delta_p)] \\ f_i^+ &= e^{c\delta_i} (x_i - y_i) \\ &= \delta_i^{-1} e^{c(\delta_i + \delta_p - \gamma_i - \gamma_p)} \sinh c(\gamma_i - \gamma_p) [V_{ip} \delta_p - F_{ip} / (\delta_i - \delta_p)] \end{aligned} \quad (3.3-12)$$

where we have used the expressions (2.3-4) and (2.3-5) for f^- and f^+ .

When $i = p$,

$$f_p^- = e^{c\delta_p} (x_p + y_p) - e^{2c\delta_p} = -e^{2c(\delta_p - \gamma_p)} [A_1 (\sinh 2c\gamma_p) / 2 + A_2] \quad (3.3-13)$$

where

$$\begin{aligned} A_1 &= 2v_{pp} + (\gamma_p^2 - \delta_p^2) \delta_p^{-2} - \sum'_m \left[V_{pm} V_{mp} + \frac{V_{pm} F_{mp} + V_{mp} F_{pm}}{\delta_m^2 - \delta_p^2} \right] \\ A_2 &= \sum'_m \frac{(\cosh 2c\gamma_p - e^{-2c\gamma_m})}{2\delta_m \delta_p (\delta_m^2 - \delta_p^2)} [V_{pm} F_{mp} \delta_m^2 + V_{mp} F_{pm} \delta_p^2 + F_{pm} F_{mp}] \end{aligned}$$

The expression for f_p^+ may be obtained in the same way but it is slightly more complicated.

$$f_p^+ = e^{c\delta_p}(x_p - y_p) = e^{2c(\delta_p - \gamma_p)}[1 - A_3 + A_4(\sinh 2c\gamma_p)/2 + A_5] \quad (3.3-14)$$

where

$$A_3 = (1 - e^{-4c\gamma_p})(\gamma_p - \delta_p)^2/(4\delta_p^2)$$

$$A_4 = \sum_m' e^{-2c\gamma_m} \left[-V_{pm} V_{mp} + \frac{V_{pm} F_{mp} - V_{mp} F_{pm}}{\delta_m^2 - \delta_p^2} + \frac{2F_{pm} F_{mp}}{(\delta_m^2 - \delta_p^2)^2} \right]$$

$$A_5 = \sum_m' \frac{(e^{-c\gamma_m} \cosh 2c\gamma_p - 1)}{2\delta_m \delta_p (\delta_m^2 - \delta_p^2)} \cdot \left[V_{pm} F_{mp} \delta_m^2 - V_{mp} F_{pm} \delta_p^2 + \frac{(\delta_m^2 + \delta_p^2) F_{pm} F_{mp}}{\delta_m^2 - \delta_p^2} \right]$$

There are several points we should mention about these formulas for f_p^- and f_p^+ : The summations with respect to m run from 1 to ∞ with the term $m = p$ omitted. γ_j and δ_j are the propagation constants of the j th mode in the bend and in the straight portion, respectively. The difference $\gamma_p^2 - \delta_p^2$ may be expressed in terms of the F 's by equation (3.2-2). In the course of obtaining (3.3-13) and (3.3-14) relations of the following sort were used.

$$t_p(1 + t_p)^{-2} = e^{-2c\gamma_p}(\sinh 2c\gamma_p)/2$$

$$(t_m + t_p^2)(1 + t_p)^{-2}(1 + t_m)^{-1} = e^{-2c\gamma_p}(\cosh 2c\gamma_p - e^{-2c\gamma_m})$$

The term A_3 arises when we subtract $(1 - t_p^2)(1 + t_p)^{-2}$ from

$$\delta_p/(\delta_p + \gamma_p t_p) - \delta_p t_p/(\gamma_p + \delta_p t_p)$$

Since $\gamma_p - \delta_p$ is $O(\xi^2)$ for a circular bend in a rectangular guide $(\gamma_p - \delta_p)^2$ is $O(\xi^4)$ and hence A_3 is negligible in the cases we shall consider.

The reflected wave set up by an incident wave of unit amplitude and containing only the p^{th} mode (i.e. the incident wave described at the beginning of this section) is given by the column matrix f^- whose elements may be obtained from (3.3-12) and (3.3-13). Likewise, the transmitted wave is given by f^+ .

PART IV

GENTLE CIRCULAR BENDS IN RECTANGULAR WAVE GUIDES

4.1 Propagation of Dominant Mode in a Gentle Bend— H in Plane of Bend

When the magnetic intensity H lies in the plane of the bend, $H_y = 0$, and equations (1.2-5) show that $B = 0$. Thus we have to deal only with

A. In order to study the dominant mode we set $\ell = 0$ in the $\cos(\pi \ell y/b)$ (A depends on y through this factor) in the formulas of Section 1.3 which involve A and assume the dimensions of the guide to be such that $b < a$.

We wish to determine γ_1^2 , the first latent root of Γ_α^2 defined by (1.3-5), from the approximate formula (3.2-2). In our case the elements δ_m^2 of diagonal matrix Γ_0^2 are obtained by putting $n(=\ell)$ to zero in (1.1-5):

$$\delta_m^2 = \Gamma_{m0}^2 = \sigma^2 + (\pi m/a)^2, \quad m = 1, 2, 3, \dots \quad (4.1-1)$$

so that (3.2-2) becomes

$$\gamma_1^2 = \Gamma_{10}^2 + F_{11} - \sum_{m=2}^{\infty} F_{1m} F_{m1} a^2 \pi^{-2} (m^2 - 1)^{-1} \quad (4.1-2)$$

The first task is to find the elements of the matrix F where, from (3.1-1) and (1.3-5),

$$F = \Gamma_\alpha^2 - \Gamma_0^2 = (P^{-1} - I)\Gamma_0^2 + P^{-1}S \quad (4.1-3)$$

In the case under consideration $P = I + R$ where R is a square matrix whose elements are very small. In fact, the asymptotic expressions leading to (A1-18) show that R_{ii} and S_{ii} are $O(\xi^2)$, with $\xi = a/\rho_1$, while R_{ij} and S_{ij} are $O(\xi)$ if $i + j$ is odd and $O(\xi^2)$ if $i + j$ is even. When the approximate value of P^{-1} obtained from (A2-2) is set in (4.1-3) and the matrix multiplications carried out it is found that

$$\begin{aligned} F_{ij} &= -R_{ij}\Gamma_{j0}^2 + S_{ij} + O(\xi^2) \\ F_{ii} &= \left(-R_{ii} + \sum_{m=1}^{\infty} R_{im} R_{mi}\right) \Gamma_{i0}^2 + S_{ii} - \sum_{m=1}^{\infty} R_{im} S_{mi} + O(\xi^3) \end{aligned} \quad (4.1-4)$$

The "order of" symbol $O(\)$ will be omitted in the following equations, it being understood that the terms in the principal diagonal are correct to within $O(\xi^2)$ and the others to within $O(\xi)$.

The values of the F 's which enter (4.1-2) may be computed from the asymptotic expressions (A1-18) for the R 's and S 's. They turn out to be

$$\begin{aligned} F_{1m} &= -4\xi m[4\Gamma_{10}^2 \pi^{-2} (m^2 - 1)^{-2} + 3a^{-2} (m^2 - 1)^{-1}] \\ F_{m1} &= -4\xi m[4\Gamma_{10}^2 \pi^{-2} (m^2 - 1)^{-2} + a^{-2} (m^2 - 1)^{-1}] \\ F_{11} &= \xi^2 [\Gamma_{10}^2 (1 - 6\pi^{-2}) + 6a^{-2}]/12 \end{aligned} \quad (4.1-5)$$

In the expressions for F_{1m} and F_{m1} m is supposed to have the values 2, 4, 6, $\dots$. For odd values of m F_{1m} and F_{m1} are $O(\xi^2)$. When $i = 1$ in the expression (4.1-4) for F_{11} , the two series therein reduce to S_3 and S_4 where

$$S_p = \sum_{m=2,4,6,\dots} m^2 (m^2 - 1)^{-p} \quad (4.1-6)$$

By expanding the typical terms in partial fractions and using the fact that the sums of (see, for example, page 238 of Reference¹⁰)

$$U_q = 1 + 3^{-q} + 5^{-q} + \dots \quad (4.1-7)$$

$$= (-)^{q/2-1} \frac{1}{2} (2^q - 1) B_q \pi^q / q!$$

for $q = 2, 4, 6$, are $\pi^2/8$, $\pi^4/96$, $\pi^6/960$, it may be shown that

$$S_3 = \pi^2/64, \quad S_4 = \pi^4/768 - \pi^2/128, \quad (4.1-8)$$

$$S_5 = (15\pi^2 - \pi^4)/3072.$$

In (4.1-7) B_q denotes the q^{th} Bernoulli number. The values of S_p may also be computed in succession from the two relations*

$$U_{2p} = \sum_{i=1}^p 2^{2i-1} C_{p+i-1, 2i-1} S_{p+i} = \sum_{i=1}^{p+1} 2^{2i-2} C_{p+i-1, 2i-2} S_{p+i}$$

where $C_{m,n}$ is a binomial coefficient. Still another method is to make use of the generating function

$$\sum_{p=0}^{\infty} t^p S_{p+1} = (1+t) \sum_{m=2,4,6,\dots} (m^2 - 1 - t)^{-1} = \frac{1}{2} - \frac{1}{2} \pi x \cot \pi x$$

where $4x^2 = 1 + t$. Note that by this definition S_1 is $\frac{1}{2}$ in contrast to the non-convergent series obtained by putting $p = 1$ in (4.1-6).

Substituting the values for the F 's given by (4.1-5) in the expression (4.1-2) for γ_1^2 and using the sums (4.1-8) of the series which occur gives

$$\gamma_1^2 = \Gamma_{10}^2 - \frac{\xi^2}{4a^2} [1 + a^2 \Gamma_{10}^2 (1 - 6\pi^{-2}) + (a\Gamma_{10}/\pi)^4 (5 - \pi^2/3)] \quad (4.1-9)$$

When the dominant mode is propagated without attenuation both γ_1^2 and Γ_{10}^2 are negative.

The general form of (4.1-9) has been obtained by both Buchholz¹ and Marshak⁵ by different methods. In our notation their result is

$$\gamma_{mn}^2 = \Gamma_{mn}^2 - \frac{\xi^2}{4a^2} \left[1 + a^2 \Gamma_{mn}^2 (1 - 6\pi^{-2} m^{-2}) + \left(\frac{a\Gamma_{mn}}{\pi m} \right)^4 (5 - \pi^2 m^2/3) \right] \quad (4.1-10)$$

where γ_{mn} is the propagation constant for the m, n^{th} mode when the magnetic vector is in the plane of the bend.

* I am indebted to John Riordan for these relations.

4.2 Reflection Due to Dominant Mode Incident upon Gentle Bend—H in Plane of Bend

Let the system be the one described in the first paragraph of Section 2.3 and let the incident wave contain only the dominant mode. Then the matrix propagation constant is the Γ_α of Section 4.1 and the column matrix h specifying the incident wave has unity for its top element and zero for its remaining elements, i.e., $p = 1$ in the formulas of Section 3.3.

We shall be interested only in the reflection coefficient, f_1^- of the dominant mode. Here we shall denote it by g_{10}^- , in line with the notation of equation (1.1-3), in order to distinguish it from the corresponding coefficient (which will be denoted by d_{01}^-) when E lies in the plane of the bend.

Setting $p = 1$ in the expression (3.3-13) for the reflection coefficient and using equation (4.1-1) for δ_m gives

$$f_1^- = g_{10}^- = -e^{2c(\Gamma_{10} - \gamma_1)} [A_1 (\sinh 2c\gamma_1)/2 + A_2] \quad (4.2-1)$$

where γ_1 has just been obtained in (4.1-9) and

$$\begin{aligned} A_1 = & 2v_{11} + (\gamma_1^2 - \Gamma_{10}^2)\Gamma_{10}^{-2} \\ & - \sum_{m=2}^{\infty} \left[V_{1m} V_{m1} + \frac{V_{1m} F_{m1} + V_{m1} F_{1m}}{\pi^2 a^{-2}(m^2 - 1)} \right], \\ A_2 = & \sum_{m=2}^{\infty} \frac{\cosh 2c\gamma_1 - e^{-2c\gamma_m}}{2\Gamma_{m0}\Gamma_{10}\pi^2 a^{-2}(m^2 - 1)} \\ & \cdot [V_{1m} F_{m1}\Gamma_{m0}^2 + V_{m1} F_{1m}\Gamma_{10}^2 + F_{1m} F_{m1}] \end{aligned} \quad (4.2-2)$$

From (A1-18) and $V_{11} = 1 + v_{11}$ it follows that

$$\begin{aligned} v_{11} &= \xi^2(1 - 6\pi^{-2})/12 \\ V_{1m} &= V_{m1} = 8\pi^{-2}\xi m(m^2 - 1)^{-2} \end{aligned} \quad (4.2-3)$$

where $m = 2, 4, 6, \dots$. For odd values of m , V_{1m} and V_{m1} are $O(\xi^2)$. Substituting these values together with those for the F 's given by (4.1-5), using the sums (4.1-8) and the expression (4.1-9) for $\gamma_1^2 - \Gamma_{10}^2$ finally leads to (after considerable cancellation)

$$A_1 = -\xi^2\Gamma_{10}^{-2}a^{-2}/4 \quad (4.2-4)$$

Likewise, for even values of m ,

$$V_{1m}F_{m1}\Gamma_{m0}^2 + V_{m1}F_{1m}\Gamma_{10}^2 + F_{1m}F_{m1} = 16\xi^2m^2a^{-4}(m^2 - 1)^{-2} \quad (4.2-5)$$

All of the terms in the expression (4.2-1) for g_{10}^- are now known (the values of γ_m may be obtained by setting $n = 0$ in (4.1-10)). We shall make the

further approximation of putting Γ_{m0} for γ_m . Since $\Gamma_{m0} - \gamma_m$ is $O(\xi^2)$ no serious error is introduced and we have

$$g_{10}^- = \frac{\xi^2 \sinh 2c\Gamma_{10}}{8\Gamma_{10}^2 a^2} - \frac{\xi^2 8}{\pi^2} \sum_{m=2,4,6,\dots} \frac{(\cosh 2c\Gamma_{10} - e^{-2c\Gamma_{m0}})}{\Gamma_{10} \Gamma_{m0} a^2} \frac{m^2}{(m^2 - 1)^3} \quad (4.2-6)$$

in which

$$a\Gamma_{m0} = [\pi^2(m^2 - 1) + a^2\Gamma_{10}^2]^{\frac{1}{2}}, \quad \xi = a/\rho_1. \quad (4.2-7)$$

For frequencies such that only the dominant mode is propagated the ratio of the power in the reflected wave to the power in the incident wave is $|g_{10}^-|^2$. Marshak has given an expression for this ratio which is the same as that obtained from (4.2-6) when the negligible (for his case) terms $e^{-2c\Gamma_{m0}}$ are omitted.

The corresponding expression for the transmission coefficient derived from (3.3-14) for f_1^+ is not as simple as (4.2-6).

4.3 Propagation of Dominant Mode in a Gentle Bend— E in Plane of Bend

When the electric intensity E lies in the plane of the bend, $E_y = 0$, and equations (1.2-5) show that $A = 0$. Here we deal with B in much the same way as we dealt with A in Section 4.1. The dominant mode is obtained by setting $\ell = 1$ in the $\sin(\pi\ell y/b)$ in the formulas pertaining to B in Section 1.3. It is assumed that $b > a$.

Examination of the matrices (1.3-14) indicates that, for the sake of convenience, we should call the top row of our matrices the 0th row and the left-most column the 0th column. In line with this we call γ_0 the propagation constant of the dominant mode in the bend. The elements δ_m^2 of the diagonal matrix Γ_0^2 are obtained by putting $n(= \ell) = 1$ in (1.1-5):

$$\delta_m^2 = \Gamma_{m1}^2 = \sigma^2 + (\pi m/a)^2 + \pi^2/b^2, \quad m = 0, 1, 2, \dots \quad (4.3-1)$$

When we make the appropriate shift in the subscripts, equation (3.2-2) yields

$$\gamma_0^2 = \Gamma_{01}^2 + F_{00} + \sum_{m=1}^{\infty} F_{0m} F_{m0} a^2 \pi^{-2} m^{-2} \quad (4.3-2)$$

in which the elements of the matrix F are to be determined from (1.3-13):

$$F = \Gamma_\beta^2 - \Gamma_0^2 = (Q^{-1} - I)\Gamma_0^2 + Q^{-1}U \quad (4.3-3)$$

As in (4.1-4) we have, with $Q = I + T$,

$$F_{ij} = -T_{ij}\Gamma_{j1}^2 + U_{ij}$$

$$F_{ii} = \left(-T_{ii} + \sum_{m=0}^{\infty} T_{im} T_{mi}\right) \Gamma_{i1}^2 + U_{ii} - \sum_{m=0}^{\infty} T_{im} U_{mi}. \quad (4.3-4)$$

By using the asymptotic expressions (A1-19) for the T 's and U 's and summing the series with the value of (4.1-7) for $q = 4$ given in Section 4.1 we obtain

$$\begin{aligned} F_{0m} &= -2\xi[2\Gamma_{01}^2\pi^{-2}m^{-2} + a^{-2}] \\ F_{m0} &= -8\xi\Gamma_{01}^2\pi^{-2}m^{-2} \\ F_{00} &= \xi^2\Gamma_{01}^2/12 \end{aligned} \quad (4.3-5)$$

In these expressions m is supposed to have the values 1, 3, 5, $\dots$. When m is even F_{0m} and F_{m0} are $O(\xi^2)$.

Substituting (4.3-5) in (4.3-2) and summing the series with the help of the values of (4.1-7) given in Section 4.1 gives

$$\gamma_0^2 = \Gamma_{01}^2 - \xi^2\Gamma_{01}^2(5 + 2a^2\Gamma_{01}^2)/60 \quad (4.3-6)$$

A result equivalent to (4.3-6) has been given by Buchholz who also gives the approximation to the propagation constant when $m > 0$ (and the electric vector in the plane of the bend). In our notation his approximation is

$$\begin{aligned} \gamma_{mn}^2 = \Gamma_{mn}^2 + \frac{\xi^2}{4a^2} \left[3 - \left(\frac{a\Gamma_{mn}}{\pi m} \right)^2 (10 + \pi^2 m^2) \right. \\ \left. + \frac{1}{3} \left(\frac{a\Gamma_{mn}}{\pi m} \right)^4 (21 + \pi^2 m^2) \right] \end{aligned} \quad (4.3-7)$$

In writing (4.3-7) we have corrected a misprint in Buchholz's expression. In order to agree with Buchholz's equation (5.30a) the leading term within the square brackets would have to be changed from 3 to -3 . This change was indicated by the results obtained when our equation (3.2-2) was used to obtain special cases of (4.3-7). Probably the best way of obtaining (4.3-7) is furnished by Marshak's method (WKB approximation, out to second order terms, applied to Bessel's differential equation). If one wishes to verify (4.3-7) by using Marshak's report⁵ as a guide, he should correct the misprint in Marshak's equation (12a).

4.4 Reflection Due to Dominant Mode Incident upon Gentle Bend— E in Plane of Bend

The problem here is the same as that treated in Section 4.2 except that now the electric vector lies in the plane of the bend. In line with equation (1.1-4), the reflection coefficient f_1^- for the dominant mode will be denoted by $\bar{d}_{01}$. As in Section 4.3 the subscripts indicating the position of matrix elements will be adjusted so as to start with 0 instead of 1. The square matrix W given by (1.4-5), and associated with the junction conditions for

B in the same manner as V is associated with A , now replaces V . Thus our expression (3.3-13) for the reflection coefficient becomes

$$\bar{f}_{01} = \bar{d}_1 = -[A_1(\sinh 2c\Gamma_{01})/2 + A_2] \quad (4.4-1)$$

where we have neglected the difference in Γ_{01} and γ_0 and where

$$A_1 = 2w_{00} + (\gamma_0^2 - \Gamma_{01}^2)\Gamma_{01}^{-2} - \sum_{m=1}^{\infty} \left[W_{0m}W_{m0} + \frac{W_{0m}F_{m0} + W_{m0}F_{0m}}{\pi^2 a^{-2}m^2} \right] \quad (4.4-2)$$

$$A_2 = \sum_{m=1}^{\infty} \frac{\cosh 2c\Gamma_{01} - e^{-2c\Gamma_{m1}}}{2\Gamma_{m1}\Gamma_{01}\pi^2 a^{-2}m^2} \cdot [W_{0m}F_{m0}\Gamma_{m1}^2 + W_{m0}F_{0m}\Gamma_{01}^2 + F_{0m}F_{m0}]$$

From $W_{00} = 1 + w_{00}$ and the asymptotic expressions (A1-19) it follows that, for $m = 1, 3, 5, \dots$

$$w_{00} = \xi^2/12 \quad (4.4-3)$$

$$W_{0m} = 2\xi m^{-2}\pi^{-2}, \quad W_{m0} = 4\xi m^{-2}\pi^{-2}$$

For even values of m , W_{0m} and W_{m0} are $O(\xi^2)$. Substitution of these values together with those for the F 's given by (4.3-5), using the sums (4.1-7) and expression (4.3-6) for $\gamma_0^2 - \Gamma_{01}^2$ leads to

$$A_1 = \xi^2/12$$

$$W_{0m}F_{m0}\Gamma_{m1}^2 + W_{m0}F_{0m}\Gamma_{01}^2 + F_{0m}F_{m0} = -8\xi^2\Gamma_{01}^2a^{-2}\pi^{-2}m^{-2}$$

for m odd.

Thus the reflection coefficient for the dominant mode when E lies in the plane of a gentle bend of length $2c$ is approximately

$$\bar{d}_{01} = -\frac{\xi^2 \sinh 2c\Gamma_{01}}{24} + \frac{\xi^2 4\Gamma_{01}}{\pi^4} \sum_{m=1,3,5,\dots}^{\infty} \frac{\cosh 2c\Gamma_{01} - e^{-2c\Gamma_{m1}}}{m^4\Gamma_{m1}} \quad (4.4-4)$$

where Γ_{m1} is given by (4.3-1) and $b > a$.

PART V

NUMERICAL CALCULATIONS

5.1 Bend in Plane of Magnetic Vector

Let $a/b = 2.25$ and $\lambda_0/a = 1.400$ where λ_0 is the free-space wavelength of the dominant wave striking the bend. The propagation constant

Γ_{10} of the dominant wave is obtained by setting $m = 1$, $n = 0$ in (1.1-5). The Γ 's corresponding to the higher modes may be obtained from the same formula:

$$\begin{aligned} a^2 \Gamma_{10}^2 &= -\left(\frac{2\pi a}{\lambda_0}\right)^2 + \pi^2 = -10.272 & a\Gamma_{10} &= i 3.205 \\ a^2 \Gamma_{20}^2 &= a^2 \Gamma_{10}^2 + 3\pi^2 = 19.336 & a\Gamma_{20} &= 4.397 \\ a^2 \Gamma_{30}^2 &= a^2 \Gamma_{10}^2 + 8\pi^2 = 68.684 & a\Gamma_{30} &= 8.288 \end{aligned} \quad (5.1-1)$$

We shall consider a 90° bend. The approximation (4.2-6) appropriate to gentle bends becomes

$$g_{10}^- = i\xi^2 [- .0122 \sin (5.03/\xi) + .0087 \cos (5.03/\xi)] \quad (5.1-2)$$

where the exponential terms have been omitted since they are generally negligible. In (5.1-2), $\xi = a/\rho_1$ and the arguments of the sine and cosine terms arise from $2c\Gamma_{10} = \pi a\Gamma_{10}/(2\xi)$. From (4.1-9) the approximate change in the propagation constant produced by the curvature is obtainable from

$$\gamma_1^2 - \Gamma_{10}^2 = .294\xi^2/a^2 \quad (5.1-3)$$

where γ_1 is the propagation constant of the dominant mode in the bend.

The determination of $g_{10}^\pm$ by matrix methods will be illustrated for a 90° bend in which $\rho_1/a = 0.6$. This makes $c/a = \rho_1\pi/(4a) = .4712$, $c\Gamma_{10} = i1.510$ and the appropriate equations in (2.3-5) and (2.3-4) become, upon setting $f_1^- = g_{10}^-$ and $f_1^+ = g_{10}^+$,

$$\begin{aligned} g_{10}^+ &= e^{c\Gamma_{10}}(x_1 - y_1) = (.061 + i.998)(x_1 - y_1) \\ g_{10}^- &= e^{c\Gamma_{10}}(x_1 + y_1) - e^{2c\Gamma_{10}} = (.061 + i.998)(x_1 + y_1) \\ &\quad + .993 - i.121 \end{aligned} \quad (5.1-4)$$

Here x_1, y_1 are the top elements in the column matrices x, y . The problem is to compute x and y from the matrix equations (2.3-3) with Γ replaced by $\Gamma_\alpha, \Gamma_\theta$ defined by (1.3-2) with $\ell = 0$, and h a column matrix whose elements are zero except the top one which is unity. Since the order of the matrices is infinite, an exact solution calls for an infinite amount of work. A compromise must be made between the accuracy desired and amount of work one is willing to do. The following numerical work uses third order matrices.

The first step is to compute the square matrix, obtained from (1.3-5),

$$a^2 \Gamma_\alpha^2 = P^{-1}(a^2 \Gamma_0^2 + a^2 S) \quad (5.1-5)$$

The elements of the diagonal matrix $a^2\Gamma_0^2$ are given by (5.1-1) and those of P and S by the equations and tables of Appendix I.

$$a^2\Gamma_\alpha^2 = \begin{bmatrix} 1.4429 & .8812 & .6821 \\ .8812 & 2.1250 & 1.3745 \\ .6821 & 1.3745 & 2.4879 \end{bmatrix}^{-1} \begin{bmatrix} -12.292 & 3.3447 & 1.6087 \\ -6.3039 & 16.369 & 6.9738 \\ -3.5034 & -11.3031 & 65.213 \end{bmatrix} \quad (5.1-6)$$

$$= \begin{bmatrix} -9.086 & -1.785 & -5.178 \\ .157 & 17.218 & -19.362 \\ .996 & -13.566 & 38.329 \end{bmatrix}$$

The next step is to use (5.1-6) to evaluate the coefficients of x and y in (2.3-3). The square matrices $\Gamma_{ac} \tanh \Gamma_{ac}$ and $\Gamma_{ac} \coth \Gamma_{ac}$ cause most of the computational difficulties. We shall deal with these matrices by using Sylvester's theorem (an account of this theorem is given in Section 3.9 of Reference⁹). This requires the determination of the latent roots and modal rows of $a^2\Gamma_\alpha^2$. However, it is interesting to note that the matrices in question may also be computed from $c^2\Gamma_\alpha^2$ (which is easily obtained from $a^2\Gamma_\alpha^2$) by processes which employ only matrix multiplication, addition, and inversion.

Thus, setting A^2 for $c^2\Gamma_\alpha^2$,

$$A^{-1} \sinh A = I + \frac{A^2}{3!} + \frac{A^4}{5!} + \dots$$

$$\cosh A = I + \frac{A^2}{2!} + \frac{A^4}{4!} + \dots$$

$$A \coth A = (\cosh A)(A^{-1} \sinh A)^{-1}$$

$$A \tanh A = A^2(A \coth A)^{-1}.$$

Although the series always converge, they do so too slowly to be of use in our computations. The same is true of the series

$$A \tanh A = \sum_{m=1}^{\infty} 8A^2 [(2m-1)^2 \pi^2 I + 4A^2]^{-1}.$$

For the matrices we shall encounter it appears best to use Sylvester's theorem even though this requires the determination of the latent roots and modal rows of $a^2\Gamma_\alpha^2$. The square matrix formed from the modal rows* will be denoted by κ .

* As has already been mentioned in the footnote associated with equation (1.3-9), we shall use the notation and theory set forth in Sections 3.5 and 3.6 of Reference⁹.

We shall use the relations*

$$\begin{aligned}\Gamma_{ac} \tanh \Gamma_{ac} &= \kappa^{-1} [\gamma c \tanh \gamma c]_d \kappa \\ \Gamma_{ac} \cosh \Gamma_{ac} &= \kappa^{-1} [\gamma c \coth \gamma c]_d \kappa\end{aligned}\quad (5.1-7)$$

where the subscript d on the brackets stands for "diagonal" matrix, the i^{th} element in the principal diagonal of $[\gamma c \tanh \gamma c]_d$ being $\gamma_i c \tanh \gamma_i c$ where

$$\gamma_i c = c \lambda_i^{1/2} / a \quad (5.1-8)$$

and λ_i is the i^{th} latent root of $a^2 \Gamma_\alpha^2$. In our applications γ_i is either positive real or positive imaginary.

From (5.1-6) the λ_i 's are the roots of

$$\begin{vmatrix} \lambda + 9.086 & 1.785 & 5.178 \\ -.157 & \lambda - 17.218 & 19.362 \\ -.996 & 13.566 & \lambda - 38.329 \end{vmatrix} \quad (5.1-9)$$

$$= \lambda^3 - 46.461 \lambda^2 - 101.96 \lambda + 3464.5 = 0$$

and have the values

$$\lambda_1 = -8.886, \quad \lambda_2 = 8.284, \quad \lambda_3 = 47.06 \quad (5.1-10)$$

The elements κ_{21} , κ_{31} of the modal row $[1, \kappa_{21}, \kappa_{31}]$ corresponding to λ_1 may be obtained by solving the two equations derived from the last two elements of

$$[1, \kappa_{21}, \kappa_{31}](\lambda_1 I - a^2 \Gamma_\alpha^2) = 0 \quad (5.1-11)$$

namely,

$$1.785 + (\lambda_1 - 17.218) \kappa_{21} + 13.566 \kappa_{31} = 0$$

$$5.178 + 19.362 \kappa_{21} + (\lambda_1 - 38.329) \kappa_{31} = 0$$

When the value of λ_1 from (5.1-10) is used these equations yield

$$\kappa_{21} = .1593, \quad \kappa_{31} = .1750$$

Likewise, the first and third elements of

$$[\kappa_{12}, 1, \kappa_{32}](\lambda_2 I - a^2 \Gamma_\alpha^2) = 0$$

and the first and second elements of

$$[\kappa_{13}, \kappa_{23}, 1](\lambda_3 I - a^2 \Gamma_\alpha^2) = 0$$

* This is the modal row matrix analogue of equation (11) in Section 3.6 of the Reference⁹. The modal rows of Γ_α are equal to the modal rows of $a^2 \Gamma_\alpha^2$.

give

$$\kappa_{12} = .0465 \quad \kappa_{32} = .6524$$

$$\kappa_{13} = .0165 \quad \kappa_{23} = -.4555$$

Thus, the numbers entering (5.1-7) are

$$\gamma_{1C} = .4712 (-8.886)^{1/2} = i 1.404, \quad \gamma_{2C} = .4712 (8.284)^{1/2} = 1.356$$

$$\gamma_{1C} \tanh \gamma_{1C} = -8.382$$

$$\gamma_{2C} \tanh \gamma_{2C} = 1.187$$

$$\gamma_{1C} \coth \gamma_{1C} = .2354$$

$$\gamma_{2C} \coth \gamma_{2C} = 1.549$$

$$\gamma_{3C} = .4712 (47.06)^{1/2} = 3.233$$

$$\gamma_{3C} \tanh \gamma_{3C} = 3.228$$

$$\gamma_{3C} \coth \gamma_{3C} = 3.243$$

$$\kappa = \begin{bmatrix} 1 & .1593 & .1750 \\ .0465 & 1 & .6524 \\ .0165 & -.4555 & 1 \end{bmatrix}$$

For the purpose of calculation it is convenient to transform (2.3-3) by inserting (5.1-7) and premultiplying by κV^{-1} . We obtain

$$\begin{aligned} ([\gamma_C \tanh \gamma_C]_d \kappa + \kappa V^{-1} \Gamma_{0C}) x &= \kappa V^{-1} c \Gamma_{0e} e^{\Gamma_0} h \\ ([\gamma_C \coth \gamma_C]_d \kappa + \kappa V^{-1} \Gamma_{0C}) y &= \kappa V^{-1} c \Gamma_{0e} e^{\Gamma_0} h \end{aligned} \quad (5.1-12)$$

in which

$$\begin{aligned} \kappa V^{-1} &= \begin{bmatrix} 1 & .1593 & .1750 \\ .0465 & 1 & .6524 \\ .0165 & -.4555 & 1 \end{bmatrix} \begin{bmatrix} 1.1204 & .3911 & .1629 \\ .3911 & 1.2833 & .4946 \\ .1629 & .4946 & 1.3460 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} .9492 & -.1992 & .0883 \\ -.2608 & .7686 & .2339 \\ .1427 & -.7900 & 1.0160 \end{bmatrix}, \quad \begin{aligned} \Gamma_{10C} &= i 1.510 \\ \Gamma_{20C} &= 2.072 \\ \Gamma_{30C} &= 3.905 \end{aligned} \end{aligned}$$

where the elements of V are obtained from the formulas and tables of Appendix I.

The i^{th} equation of the set obtained by writing out the first of equations (5.1-12) is

$$\sum_{j=1}^3 [\kappa_{ji} \gamma_{jC} \tanh \gamma_{jC} + (\kappa V^{-1})_{ij} \Gamma_{j0C}] x_j = (\kappa V^{-1})_{i1} c \Gamma_{01} e^{\Gamma_{01}} \quad (5.1-13)$$

where $(\kappa V^{-1})_{ij}$ denotes the element in the i^{th} row and j^{th} column of κV^{-1} , κ_{ji} is the element in the i^{th} row and j^{th} column (note the reversal of the usual convention regarding the order of subscripts) of κ , $\kappa_{ii} = 1$, and h has disappeared because it is a column matrix whose top element is unity while the remaining elements are zero. It will be noted that the only

imaginary terms in (5.1-13) occur in the coefficients of x_0 and arise from the imaginary quantity $\Gamma_{10}c$. By making the substitution

$$x_j = \frac{u_j c \Gamma_{10} e^{c \Gamma_{10}}}{1 + u_1 c \Gamma_{10}} \quad (5.1-14)$$

the set (5.1-13) may be reduced to

$$(\kappa_1 \gamma_i c \tanh \gamma_i c) u_1 + \sum_{j=2}^3 [\kappa_j \gamma_i c \tanh \gamma_i c + (\kappa V^{-1})_{ij} \Gamma_{j0} c] u_j = (\kappa V^{-1})_{i1} \quad (5.1-15)$$

in which the coefficients are all real. It should be noticed, however, that nothing is gained by making the substitution (5.1-14) when the frequency is so high that other modes in addition to the dominant are propagated.

The equation for y corresponding to (5.1-15) may be obtained by replacing $\tanh$ by $\coth$ and u by v where now

$$y_j = \frac{v_j c \Gamma_{10} e^{c \Gamma_{10}}}{1 + v_1 c \Gamma_{10}} \quad (5.1-16)$$

Incidentally, if we set $j = 1$ in (5.1-14) and (5.1-16) and substitute in the expressions (5.1-4) for $g_{10}^{\pm}$ we may show that, since u_1 and v_1 are real,

$$|g_{10}^+|^2 + |g_{10}^-|^2 = 1 \quad (5.1-17)$$

Equation (5.1-17) may be obtained at once from the fact that the energy of the waves leaving the bend must equal the energy of the incident wave. It may also be shown that g_{10}^- vanishes when $u_1 v_1 c^2 \Gamma_{10}^2 = 1$.

When the above numbers are set in the three equations obtained from (5.1-15) we get

$$\begin{aligned} -8.382 u_1 - 1.748 u_2 - 1.122 u_3 &= .9492 \\ .055 u_1 + 2.780 u_2 + 1.688 u_3 &= -.2608 \\ .053 u_1 - 3.105 u_2 + 7.191 u_3 &= .1427 \end{aligned}$$

from which

$$u_1 = -.0940, \quad x_1 = .1400 + i.0113$$

The equations for v_1 obtained by substituting $\coth$ for $\tanh$ are

$$\begin{aligned} .2354 v_1 - .3732 v_2 + .3861 v_3 &= .9492 \\ .0717 v_1 + 3.1417 v_2 + 1.924 v_3 &= -.2608 \\ .0534 v_1 - 3.1143 v_2 + 7.211 v_3 &= .1427 \end{aligned}$$

from which

$$v_1 = 3.930, \quad y_1 = -.1045 + i.9803$$

When these values are set in (5.1-4) we finally obtain

$$\begin{aligned}g_{10}^{+} &= .9822 + i.1858 \\g_{10}^{-} &= .0048 - i.0255\end{aligned}$$

The following table lists values of $g_{10}^{\pm}$ obtained by the methods of this section*. Here the bend is in the plane of H , $a/b = 2.25$, $\lambda_0/a = 1.400$, where λ_0 is the free space wavelength of the incident wave. ρ_1 is the radius of curvature of the axis of the guide. The smallest possible value of ρ_1/a is 0.5. The term "approx." refers to equations (5.1-2) while "1st order", "2nd order", etc. refers to the order of the matrices used in the computations. The amplitude of the reflected wave is g_{10}^{-} and the amplitude of the wave sent forward is g_{10}^{+} when the incident wave is of unit amplitude.

g_{10}^{+}					
ρ_1/a	Approx.	1st order	2nd order	3rd order	
.6		.964 +i.267	.980 +i.197	.982 +i.186	
.7		.974 +i.224	.994 +i.105	.994 +i.111	
.8		.984 +i.178	.997 +i.082	.997 +i.082	
.9		.988 +i.153	.997 +i.074	.997 +i.073	
1.0		.991 +i.135	.998 +i.066	.998 +i.066	
1.2		.994 +i.110	.998 +i.056	.998 +i.056	
1.5		.996 +i.084	.999 +i.043	.999 +i.044	
g_{10}^{-}					
.6	-i.0280	.0020 -i.0074	.0056 -i.0280	.0048 -i.0255	
.7	-i.0068	-.0005 +i.0023	.0013 -i.0131	.0007 -i.0066	
.8	+i.0062	-.0013 +i.0074	-.0003 +i.0039	-.0004 +i.0051	
.9	+i.0128	-.0014 +i.0087	-.0009 +i.0123	-.0009 +i.0123	
1.0	+i.0143	-.0010 +i.0075	-.0010 +i.0148	-.0010 +i.0147	
1.2	+i.0079	-.0002 +i.0018	-.0005 +i.0086	-.0005 +i.0085	
1.5	-i.0040	+.0003 -i.0034	+.0002 -i.0041	+.0002 -i.0042	

It appears that the values obtained from the first order matrices are quite far from the true values. On the other hand there is considerable agreement between the approximation and the second and third order values, especially at the larger values of ρ_1/a .

5.2 Bend in Plane of Electric Vector

The calculations for this case are quite similar to those presented in Section 5.1. If we are to deal with the same waveguide it is necessary to

* The computations were performed by Miss M. Darville. I am also indebted to her for the values given in the tables in Section 5.2 and Appendix I.

interchange the dimensions a and b so that now $b/a = 2.25$ and, for the same frequency, $\lambda_0/b = 1.400$.

For a 90° bend the approximation (4.4-4) for the reflection coefficient d_{01}^- for gentle bends, i.e. for $\xi = a/\rho_1$ small, becomes

$$d_{01}^- = i\xi^2[-.0417 \sin (2.23/\xi) + .0209 \cos (2.23/\xi)]$$

where the negligible exponential terms have been neglected just as in the analogue (5.1-2) for $\bar{g}_{10}$.

The following table, which is similar to the table at the end of Section 5.1, gives the results of computations for bends in the plane of the electric vector.

		d_{10}^+	
ρ_1/a	Approx.	1st order	2nd order
.6		.823 + i .547	.975 + i .223
.7		.887 + i .447	.994 + i .051
.8		.921 + i .380	.996 + i .042
.9		.941 + i .332	.997 + i .035
1.0		.954 + i .295	.998 + i .031
1.2		.970 + i .242	.999 + i .023
1.5		.982 + i .190	1.000 + i .017
		d_{10}^-	
.6	- i .0996	-.0855 + i .1284	-.0020 + i .0086
.7	- i .0848	-.0520 + i .1031	+.0050 - i .0975
.8	- i .0706	-.0330 + i .0800	+.0033 - i .0792
.9	- i .0575	-.0214 + i .0605	+.0022 - i .0635
1.0	- i .0457	-.0137 + i .0443	.0021 - i .0507
1.2	- i .0258	-.0051 + i .0204	.0007 - i .0282
1.5	- i .0051	+.0001 - i .0004	.0001 - i .0062

The agreement between the approximation for d_{10}^- and its second order matrix value is fairly good from $\rho_1/a = .7$ onward.

APPENDIX I

CALCULATION OF P_{pm} ETC. FOR CIRCULAR BEND

It is convenient to write P_{pm} and Q_{pm} as given by (1.2-10) and (1.2-15) in the form

$$\begin{aligned} P_{pm} &= \delta_m^p + R_{pm}, & p, m &= 1, 2, 3, \dots \\ Q_{km} &= \delta_m^p + T_{pm}, & p, m &= 0, 1, 2, \dots \end{aligned} \quad (\text{A1-1})$$

where δ_m^p is unity if $p = m$ and is zero otherwise and

$$R_{pm} = (2/a) \int_0^a (\rho_1^2 \rho^{-2} - 1) \sin(\pi p x/a) \sin(\pi m x/a) dx \quad (A1-2)$$

$$T_{pm} = (\epsilon_p/a) \int_0^a (\rho_1^2 \rho^{-2} - 1) \cos(\pi p x/a) \cos(\pi m x/a) dx$$

in which $\epsilon_0 = 1$, $\epsilon_p = 2$, $p = 1, 2, \dots$

In (A1-2), (1.2-11), (1.2-16), (1.4-3), (1.4-6) we make the substitutions

$$\begin{aligned} p - m &= r, & u &= \rho_1 \pi/a - \pi/2 = \rho_2 \pi/a \\ p + m &= s, & v &= \rho_1 \pi/a + \pi/2 = \rho_3 \pi/a \\ y &= \pi x/a, & w &= \rho_1 \pi/a, & \rho &= x + \rho_1 - a/2 = a(y + u)/\pi \end{aligned} \quad (A1-3)$$

Introduction of the integrals

$$\begin{aligned} I_s &= (1/\pi) \int_0^\pi [w^2(y + u)^{-2} - 1] \cos sy dy \\ &= (1/a) \int_0^a (\rho_1^2 \rho^{-2} - 1) \cos(\pi s x/a) dx \end{aligned} \quad (A1-4)$$

$$J_s = \pi \int_0^\pi \frac{\sin sy}{y + u} dy = \pi \int_0^a \sin(\pi s x/a) dx/\rho,$$

$$K_s = \frac{\rho_1}{a} \int_0^\pi \frac{\cos sy}{y + u} dy$$

enables us to write

$$\begin{aligned} R_{pm} &= I_r - I_s, & S_{pm} &= -ma^{-2}(J_s + J_r) \\ T_{pm} &= \epsilon_p(I_r + I_s)/2, & U_{pm} &= m\epsilon_p a^{-2}(J_s - J_r)/2 \\ V_{pm} &= K_r - K_s, & W_{pm} &= \epsilon_p(K_r + K_s)/2 \end{aligned} \quad (A1-5)$$

where I_s and K_s are even functions of s and J_s is an odd function of s . $\epsilon_0 = 1$ and $\epsilon_p = 2$, $p = 1, 2, 3, \dots$. Since w and u depend only upon the ratio ρ_1/a , the values of I_s , K_s and J_s depend only upon ρ_1/a and the integer s . These quantities are tabulated at the end of this appendix.

Setting $y + u$ equal to t gives

$$\begin{aligned} J_s &= \pi \int_u^v \sin s(t - u) dt/t \\ &= \pi[Si(sv) - Si(su)] \cos su - \pi[Ci(sv) - Ci(su)] \sin su \end{aligned} \quad (A1-6)$$

where Si and Ci denote the integral sine and cosine functions. Integrating by parts enables us to express I_s in terms of J_s . Thus

$$\int_0^\pi (y + u)^{-2} \cos sy dy = u^{-1} - v^{-1} \cos s\pi - \pi^{-1} s J_s \quad (A1-7)$$

and

$$I_s = \pi^{-1} w^2 [u^{-1} - (-)^s v^{-1} - \pi^{-1} s J_s] \quad (\text{A1-8})$$

except when $s = 0$ in which case

$$\begin{aligned} I_0 &= \pi^{-1} w^2 (u^{-1} - v^{-1}) - 1 = w^2 / (uv) - 1 \\ &= \pi^2 / (4uv) = [(2\rho_1/a)^2 - 1]^{-1} \end{aligned} \quad (\text{A1-9})$$

When ρ_1/a is large, u and v are large, and the asymptotic expansion of (A1-6) gives

$$\pi^{-1} J_s \sim s^{-1} [u^{-1} - (-)^s v^{-1}] - 2! s^{-3} [u^{-3} - (-)^s v^{-3}] + \dots \quad (\text{A1-10})$$

When (A1-10) is placed in (A1-8)

$$I_s \sim \pi^{-1} 2! w^2 s^{-2} [u^{-3} - (-)^s v^{-3}] - \dots \quad (\text{A1-11})$$

Formulas for K_s may be obtained in much the same way.

$$\begin{aligned} K_s &= (\rho_1/a) \int_u^v \cos s(t-u) dt/t \\ &= (\rho_1/a) \{ [Ci(sv) - Ci(su)] \cos su + [Si(sv) - Si(su)] \sin su \} \end{aligned} \quad (\text{A1-12})$$

and when $s = 0$

$$K_0 = (\rho_1/a) \log (1 + \pi/u) \quad (\text{A1-13})$$

The asymptotic expression is

$$a\rho_1^{-1} K_s \sim s^{-2} [u^{-2} - (-)^s v^{-2}] - 3! s^{-4} [u^{-4} - (-)^s v^{-4}] + \dots$$

It is convenient to write the asymptotic expressions in terms of the new variable

$$\xi = a/\rho_1 \quad (\text{A1-14})$$

When s is even and greater than zero

$$J_s \sim \xi^2/s, \quad I_s \sim 6\xi^2\pi^{-2}s^{-2}, \quad K_s \sim 2\xi^2\pi^{-2}s^{-2} \quad (\text{A1-15})$$

and when s is odd

$$J_s \sim 2\xi/s, \quad I_s \sim 4\xi\pi^{-2}s^{-2}, \quad K_s \sim 2\xi\pi^{-2}s^{-2} \quad (\text{A1-16})$$

When $s = 0$

$$I_0 \sim \xi^2/4, \quad K_0 \sim 1 + \xi^2/12 \quad (\text{A1-17})$$

We shall need the following asymptotic expressions which may be obtained from the above work

$$R_{11} \sim \frac{\xi^2}{4} \left(1 - \frac{6}{\pi^2} \right), \quad S_{11} \sim -a^{-2} \xi^2/2, \quad V_{11} = 1 + \xi^2 \left(\frac{1}{12} - \frac{1}{2\pi^2} \right)$$

$$R_{1m} = R_{m1} \sim 16\xi m\pi^{-2}(m^2 - 1)^{-2} \quad (\text{A1-18})$$

$$S_{1m} \sim -S_{m1}, \quad S_{1m} \sim 4a^{-2}\xi m(m^2 - 1)^{-1}$$

$$V_{1m} = V_{m1} \sim 8\xi m\pi^{-2}(m^2 - 1)^{-2} \sim R_{1m}/2$$

where $m = 2, 4, 6, \dots$. Also, if now $m = 1, 3, 5, \dots$,

$$T_{00} \sim \xi^2/4, \quad U_{00} = 0, \quad W_{00} = 1 + \xi^2/12$$

$$T_{m0} \sim 8\xi m^{-2}\pi^{-2}, \quad U_{m0} = 0 \quad (\text{A1-19})$$

$$T_{0m} \sim 4\xi m^{-2}\pi^{-2}, \quad U_{0m} \sim 2a^{-2}\xi$$

$$W_{0m} \sim 2\xi m^{-2}\pi^{-2}, \quad W_{m0} \sim 4\xi m^{-2}\pi^{-2}$$

Values of I_s , J_s and K_s

ρ_1/a	I_0	I_1	I_2	I_3	I_4	I_5	I_6
.5	∞	∞	∞	∞	∞	∞	∞
.6	2.2723	2.31879	1.82979	1.43755	1.14772	.94432	.78458
.7	1.04166	1.28232	.76232	.53315	.37692	.28832	.22090
.8	.64103	.88256	.44628	.29206	.18995	.14101	.09986
.9	.44643	.68200	.30111	.18992	.11546	.08517	.05908
1.0	.33333	.56052	.22008	.13637	.07818	.05814	.03865
1.1	.26041	.47844	.16916	.10459	.05772	.04298	.02759
1.2	.21008	.41905	.13486	.08413	.04344	.03361	.02056
1.3	.17361	.37361	.11059	.06961	.03500	.02732	.01633
1.4	.14620	.33780	.09259	.05926	.03168	.02293	.01297
1.5	.12500	.30876	.07872	.05147	.02315	.01969	.01068
2.0	.06667	.21803	.04133	.03080	.01135		
2.5	.04167	.16980	.02564	.02217	.00675		

	J_0	J_1	J_2	J_3	J_4	J_5	J_6
.5	0						
.6	0	3.99809	2.01979	2.31576	1.48355	1.66348	1.15716
.7	0	3.21624	1.3054	1.58176	.84936	1.04899	.61931
.8	0	2.72356	.73339	1.21541	.56683	.77644	.40134
.9	0	2.37231	.70698	.99327	.41079	.62183	.28546
1.0	0	2.10615	.55663	.84343	.31379	.52170	.21578
1.1	0	1.89624	.45091	.73508	.24849	.45123	.16981
1.2	0	1.72581	.37335	.65279	.20254	.39869	.13768
1.3	0	1.58448	.31450	.58812	.16843	.35788	.11413
1.4	0	1.46507	.26878	.53573	.14216	.32515	.09636
1.5	0	1.3628	.23251	.49238	.12243	.29825	.08254
2.0	0	1.0122	.12817	.35299	.06596		
2.5	0	.8062	.08128	.27660	.04140		

ρ_1/a	K_0	K_1	K_2	K_3	K_4	K_5	K_6
.5	∞	∞	∞	∞	∞	∞	∞
.6	1.43874	.61659	.31832	.22547	.15539	.12199	.09275
.7	1.25423	.42303	.17355	.11508	.06971	.05353	.03704
.8	1.17306	.33141	.11439	.07463	.04093	.03195	.02038
.9	1.12748	.27575	.08261	.05431	.02737	.02214	.01325
1.0	1.09861	.23761	.06305	.04244	.01976	.01675	.00938
1.1	1.07891	.20953	.04994	.03471	.01509	.01341	.00705
1.2	1.06476	.18785	.04075	.02935	.01192	.01126	.00548
1.3	1.05421	.17050	.03391	.02539	.00977	.00956	.00443
1.4	1.04610	.15626	.02871	.02243	.00807	.00838	.00365
1.5	1.03972	.14434	.02467	.02013	.00682	.00745	.00315
2.0	1.02165	.10508	.01332	.01330	.00358		
2.5	1.01366	.08295	.00838	.01010	.00217		

APPENDIX II

FUNCTIONS OF ALMOST DIAGONAL MATRICES

Let E be a matrix whose elements are small in comparison with unity. It is then often possible to approximate a matrix defined as some function of the matrix $I + E$, where I is the unit matrix, by the expansion

$$f(I + E) = If(1) + \frac{E}{1!}f'(1) + \frac{E^2}{2!}f''(1) + \dots \quad (\text{A2-1})$$

Thus, for example, when we take $f(z)$ to be z^{-1} we obtain

$$(I + E)^{-1} = I - E + E^2 - \dots \quad (\text{A2-2})$$

Here we shall give similar formal results for $f(D + E)$ where now D is a diagonal matrix

$$D = \begin{bmatrix} d_1 & 0 & \cdot & 0 \\ 0 & d_2 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & d_N \end{bmatrix} \quad (\text{A2-3})$$

whose diagonal elements are unequal and the elements E_{ij} and E_{ji} are small in comparison with the absolute value of $|d_i - d_j|$. We shall restrict ourselves to a first approximation of the non-diagonal terms of $f(D + E)$ and to a second approximation of the diagonal terms. The results are closely related to the ones obtained from the perturbation theory used in wave mechanics.

We assume that $f(D + E)$ may be defined by the series

$$f(D + E) = a_0I + a_1(D + E) + a_2(D + E)^2 + \dots \quad (\text{A2-4})$$

where a_n is a scalar and

$$(D + E)^2 = (D + E)(D + E) = D^2 + DE + ED + E^2$$

and so on. The sum of the terms independent of E is $f(D)$. The terms of order E are

$$\begin{aligned} & E \text{ in } D + E \\ & DE + ED \text{ in } (D + E)^2 \\ & D^2E + DED + ED^2 \text{ in } (D + E)^3 \\ & \Sigma D^\ell ED^m \text{ in } (D + E)^n \end{aligned} \tag{A2-5}$$

where the summation extends over the non-negative integer values of ℓ and m for which $\ell + m = n - 1$. The element in the i th row and j th column of $D^\ell ED^m$ is $d_i^\ell E_{ij} d_j^m$ and hence the corresponding element in the summation in (A2-5) is

$$E_{ij} \Sigma d_i^\ell d_j^m = \begin{cases} (d_i^n - d_j^n)/(d_i - d_j), & i \neq j \\ n d_i^{n-1}, & i = j. \end{cases} \tag{A2-6}$$

Thus the terms of order E in the i th row and j th column of $f(D + E)$ are, from (A2-6) and (A2-4),

$$\begin{aligned} & E_{ij} \frac{f(d_i) - f(d_j)}{d_i - d_j}, \quad i \neq j \\ & E_{ii} f'(d_i), \quad i = j \end{aligned} \tag{A2-7}$$

where the prime on f denotes its first derivative.

The terms of order E^2 in $(D + E)^n$ are

$$\sum_{k, \ell, m} D^k E D^\ell E D^m = \sum_{k, \ell, m} [d_i^k E_{ij}]_M [d_i^\ell E_{ij} d_j^m]_M \tag{A2-8}$$

where the summations extend over all the non-negative integer values of k, ℓ, m for which $k + \ell + m = n - 2$. On the right $[d_i^k E_{ij}]_M$ denotes a square matrix whose element in the i th row and j th column is $d_i^k E_{ij}$. Likewise the second factor in brackets is a matrix having $d_i^\ell E_{ij} d_j^m$ in the i th row and j th column. The element in the i th row and j th column of (A2-8) is, from the rule for the product of two matrices,

$$\sum_{k, \ell, m} \sum_{s=1}^N (d_i^k E_{is})(d_s^\ell E_{sj} d_j^m) = \sum_{s=1}^N E_{is} E_{sj} \sum_{k, \ell, m} d_i^k d_s^\ell d_j^m.$$

If i, s , and j are unequal the sum in k, ℓ, m is

$$\frac{1}{d_i - d_j} \left[\frac{d_i^n - d_s^n}{d_i - d_s} - \frac{d_j^n - d_s^n}{d_j - d_s} \right]$$

and in case of equality the sum may be found by a limiting process. Since we are interested in this order of approximation only for the diagonal terms we set $i = j$ and obtain for the sum

$$\frac{nd_i^{n-1}}{d_i - d_s} - \frac{d_i^n - d_s^n}{(d_i - d_s)^2}, \quad i \neq s$$

$$\frac{n(n-1)}{2!} d_i^{n-2}, \quad i = s.$$

Thus the contribution to the i th diagonal element of $f(D + E)$ from terms of type (A2-8) is

$$\frac{E_{ii}^2}{2!} f''(d_i) + \sum_{s=1}^N{}' E_{is} E_{si} \left[\frac{f'(d_i)}{d_i - d_s} - \frac{f(d_i) - f(d_s)}{(d_i - d_s)^2} \right] \quad (\text{A2-9})$$

where the prime on Σ indicates that the term $s = i$ is to be omitted.

Thus, to summarize, we may say that the first approximation to the non-diagonal term in the i th row and j th column ($i \neq j$) of $f(D + E)$ is

$$E_{ij} \frac{f(d_i) - f(d_j)}{d_i - d_j} \quad (\text{A2-10})$$

and the second approximation to the diagonal term in the i th row and i th column of $f(D + E)$ is

$$f(d_i) + E_{ii} f'(d_i) + \frac{E_{ii}^2}{2!} f''(d_i) + \sum_{s=1}^N{}' E_{is} E_{si} \left[\frac{f'(d_i)}{d_i - d_s} - \frac{f(d_i) - f(d_s)}{(d_i - d_s)^2} \right] \quad (\text{A2-11})$$

where the primes on f denote derivatives and the prime on Σ indicates that the term $s = i$ is to be omitted.

Two results obtained from (A2-10) and (A2-11) are of interest. For the first result we set $f(z) = z^{-1}$ and get the following approximations to the elements of $(D + E)^{-1}$:

$$- E_{ij} (d_i d_j)^{-1}, \quad i \neq j$$

$$d_i^{-1} - d_i^{-2} \left[E_{ii} - \sum_{s=1}^N E_{is} E_{si} d_s^{-1} \right], \quad i = j. \quad (\text{A2-12})$$

For the second result we set $f(z) = z^{1/2}$ and obtain the following approximations to the elements of $(D + E)^{1/2}$:

$$E_{ij} (d_i^{1/2} + d_j^{1/2})^{-1}, \quad i \neq j$$

$$d_i^{1/2} + \frac{1}{2} d_i^{-1/2} \left[E_{ii} - \sum_{s=1}^N E_{is} E_{si} (d_i^{1/2} + d_s^{1/2})^{-2} \right], \quad i = j. \quad (\text{A2-13})$$

In (A2-12) and (A2-13) the summations include the term $s = i$.

We shall now state several results related to the above formulas. Let u denote the matrix $D + E$ so that the typical element $u_{ij} = E_{ij}$, $i \neq j$, and $u_{ii} = d_i + E_{ii}$. Then the latent roots $\lambda_1, \lambda_2, \dots, \lambda_N$ of u are the roots of the equation obtained by setting the determinant of $\lambda I - u$ to zero:

$$|\lambda I - u| = \begin{vmatrix} \lambda - u_{11} & -u_{12} & \cdot \\ -u_{21} & \lambda - u_{22} & \cdot \\ \cdot & \cdot & \cdot \end{vmatrix} = 0. \quad (\text{A2-14})$$

The modal column k_j corresponding to the j th root λ_j satisfies the matrix equation

$$(\lambda_j I - u)k_j = 0. \quad (\text{A2-15})$$

Since the non-diagonal elements of u are small, we see from (A2-14) that we may label the roots so as to make λ_j nearly equal to u_{jj} , and this together with (A2-15) shows that all the elements of k_j are nearly zero except the j th which we may choose to be unity. When these approximate values are taken as a first approximation in the process of solving (A2-15) by successive approximations, the second approximation is found to be

$$\lambda_j = u_{jj} + \sum_{s=1}^N \frac{u_{js} u_{sj}}{u_{jj} - u_{ss}} = u_{jj} + \sum_{s=1}^N u_{js} k_{sj}$$

$$k_j = \begin{bmatrix} k_{1j} \\ k_{2j} \\ \cdot \\ 1 \\ \cdot \\ k_{Nj} \end{bmatrix} \quad k_{sj} = \frac{u_{sj}}{u_{jj} - u_{ss}}, \quad s \neq j \quad (\text{A2-16})$$

where the 1 in the column for k_j occurs as the j th element. This expression for λ_j occurs in the perturbation method often used in wave mechanics.

For the modal row κ_j corresponding to λ_j we have in much the same way

$$\kappa_j(\lambda_j I - u) = 0$$

$$\kappa_j = [\kappa_{1j}, \kappa_{2j}, \dots, \kappa_{jj}, \dots, \kappa_{Nj}] \quad (\text{A2-17})$$

$$\kappa_{sj} = \frac{u_{js} \kappa_{jj}}{u_{jj} - u_{ss}}$$

where the last expression is an approximation and where κ_{jj} may be chosen at our convenience.

The results (A2-10) and (A2-11) may also be obtained from (A2-2), (A2-16) and the relation*

$$f(u) = k \begin{bmatrix} f(\lambda_1) & 0 & \cdot & 0 \\ 0 & f(\lambda_2) & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & f(\lambda_N) \end{bmatrix} k^{-1} \quad (\text{A2-18})$$

* This is equation (12) in Section 3.6 of Reference⁹. Although proved only for polynomials it may be verified to be true for the applications which we shall make.

4. Electromagnetic Waves in a Bent Pipe of Rectangular Cross-Section, *Quart. App. Math.*, 1, 328-333 (1944). See also a note on this paper by S. A. Schelkunoff in the same journal 2, 171 (1944).
5. Theory of Circular Bends in Rectangular Wave Guides, *Rad. Lab. Report* 43-45, June 24, 1943. An abstract under the number PB2869 appears in *Bibl. Sci. and Indus. Rept.*, Vol. 1, No. 7, p. 253.
6. Electromagnetic Waves, D. Van Nostrand, New York (1943).
7. Steady State Solutions of Transmission Line Equations, *B. S. T. J.*, 20, 131-178 (1941).
8. H. T. Davis, The Theory of Linear Operators, Principia Press (1936).
9. Frazer, Duncan and Collar, Elementary Matrices, Camb. Univ. Press (1938).
10. K. Knopp, Theory and Application of Infinite Series, Blackie and Son (1928).

The Approximate Solution of Linear Differential Equations

By MARION C. GRAY and S. A. SCHELKUNOFF

Linear differential equations with variable coefficients occur in many fields of applied mathematics: in the theories of acoustics, elastic waves, electromagnetic waves in stratified media, nonuniform transmission lines, wave guides, antennas, wave mechanics. The "Wave Perturbation" method described in greater detail elsewhere¹ is particularly useful in those ranges of the independent variable in which the "WKB Approximation" is not sufficiently accurate. The present paper endeavors to illustrate the remarkable accuracy of this method, particularly when compared with Picard's method.

I. INTRODUCTION

IN A recent paper¹ the approximate solution of linear differential equations by a wave perturbation method was described. When the method was applied to equations whose exact solutions were known we were greatly impressed by the rapidity of convergence of the successive approximations. Hence the purpose of this note is to present some illustrations in the hope that others may be interested and may find the proposed method an improvement on those now in use.

In essence the wave perturbation method dates back to Liouville², but in his *mémoires* he was interested in a problem of heat conduction involving a non-homogeneous differential equation with homogeneous boundary conditions, whereas we consider a homogeneous equation

$$y'' = F(x)y \quad (1)$$

with non-homogeneous initial conditions

$$y(a) = 1, y'(a) = 0 \quad (2a)$$

or

$$y(a) = 0 \quad y'(a) = 1, \quad (2b)$$

the solution being desired in an interval $a \leq x \leq b$. Since the solution for any assigned initial or boundary conditions can be expressed as a linear combination of the solutions satisfying (2a) and (2b) we have not imposed any real limitation.

II. THEORY

Comparison of the wave perturbation method with Picard's method (which is essentially a linear perturbation method) is particularly instruc-

tive. It will be recalled that in Picard's formulation the differential equation (1) is replaced by an integral equation

$$y(x) = y(a) + (x - a)y'(a) + \int_a^x F(u)y(u)(x - u) du \quad (3)$$

where $y(a)$ and $y'(a)$ are assigned initial values*. Writing

$$L_0(x) = y(a) + (x - a)y'(a), \quad (4)$$

$$L_n(x) = \int_a^x F(u)L_{n-1}(u)(x - u) du, \quad n = 1, 2, 3, \dots,$$

the series

$$y(x) = L_0(x) + L_1(x) + L_2(x) + \dots \quad (5)$$

is shown to converge to a solution of the original equation. In practical applications, unfortunately, it is usually found that the successive approximations converge rather slowly unless the interval (a, b) is small.

In the wave perturbation method we first rewrite equation (1) in the form

$$y'' = -\beta^2 y + [\beta^2 + F(x)]y = -\beta^2 y + f(x)y, \quad (6)$$

and instead of the integral equation (3) we use

$$\begin{aligned} y(x) = & y(a) \cos \beta(x - a) + \frac{1}{\beta} y'(a) \sin \beta(x - a) \\ & + \frac{1}{\beta} \int_a^x f(u)y(u) \sin \beta(x - u) du. \end{aligned} \quad (7)$$

The parameter β is arbitrary and might be defined in various ways. We have found it convenient to use the definition

$$\beta^2 = - \frac{1}{b - a} \int_a^b F(x) dx, \quad (8)$$

so that if $F(x)$ is negative β is real and our first approximation

$$W_0(x) = y(a) \cos \beta(x - a) + \frac{1}{\beta} y'(a) \sin \beta(x - a) \quad (9)$$

is sinusoidal. If F is positive β is imaginary and we start with an exponential approximation. If F changes sign in (a, b) the best procedure is to

* This is not quite the usual form of the integral equation but it is substantially equivalent.

subdivide the interval and obtain separate approximations, though this is not necessary if F is predominantly of one sign throughout (a, b) . To (9) we now add the sequence

$$W_n(x) = \frac{1}{\beta} \int_a^x f(u) W_{n-1}(u) \sin \beta(x - u) du, \quad (10)$$

and the series

$$y(x) = W_0(x) + W_1(x) + W_2(x) + \dots \quad (11)$$

is the desired solution.

The flexibility of the wave perturbation method as compared with Picard's linear method lies essentially in the introduction of the variable parameter β . Since we make β depend on the length of the interval (a, b) in which a solution is desired the approximations may be extended over much longer intervals than is feasible in Picard's method. If $F(x)$ is a slowly varying function throughout (a, b) , so that $f(x)$ is small, it will be found that the first approximation $W_0(x)$ is good, and the second, $W_0 + W_1$ is generally adequate.

Another choice for β is

$$\beta = \frac{1}{b-a} \int_a^b \sqrt{-F(x)} dx. \quad (12)$$

However, the integration in (8) will often be simpler than in (12).

Picard's method is a special case of the wave perturbation method, with $\beta = 0$. In fact, if $F(x)$ changes sign in (a, b) , then in some cases β as defined by (8) will reduce to zero.

If $F(x)$ is a rapidly varying function, or if the solution is desired over an infinite interval, it is usually advantageous to transform equation (1) by first introducing a new independent variable

$$\theta = \int_a^x \sqrt{-F(x)} dx, \quad (13)$$

and then removing the first order term in the new equation by an appropriate transformation of the dependent variable.

III. EXAMPLES

For our illustrations we have used mainly the simple equation

$$y'' = -xy \quad (14)$$

whose exact solution can be expressed in terms of Bessel functions of order $\pm 1/3$. Since the Bessel functions are oscillatory in nature it might be

suggested that comparison with Picard's method is weighted in our favor. This does not seem to be the case, as will be illustrated in example 4 where the exact solution is a monotonically increasing function. It has also been suggested that Picard's solution might be improved by starting from a better initial approximation, say W_0 , rather than from the linear approximation L_0 , but we have not found any marked improvement in the succeeding approximations (see examples 1 and 2). The various points of interest will be brought out in our examples, with the accompanying figures, which we shall now briefly describe. In each figure the heavy curve is the accurate solution while the approximations are indicated by self-explanatory letters.

Example 1, Fig. 1

$$y'' = -xy, 0 \leq x \leq 2$$

$$y(0) = 1, y'(0) = 0$$

Exact solution: $y(x) = \Gamma(\frac{2}{3})3^{-1/3}x^{1/2}J_{-1/3}(\frac{2}{3}x^{3/2})$

(a) Wave perturbation

$$W_0 = \cos x$$

$$W_1 = -\frac{1}{4}x \cos x + \frac{1}{4}(1 + 2x - x^2) \sin x$$

(b) Linear perturbation

$$L_0 = 1$$

$$L_1 = -\frac{x^3}{6}$$

(c) Linear perturbation using initial sinusoidal approximation

$$\bar{L}_0 = \cos x = W_0$$

$$\bar{L}_1 = x + x \cos x - 2 \sin x$$

Example 2, Figs. 2, 3 and 4

$$y'' = -xy, 2 \leq x \leq 6$$

$$y(2) = 0, y'(2) = 1$$

Exact solution:

$$y(x) = -.84423x^{1/2}J_{-1/3}(\frac{2}{3}x^{3/2}) - .019291x^{1/2}J_{1/3}(\frac{2}{3}x^{3/2})$$

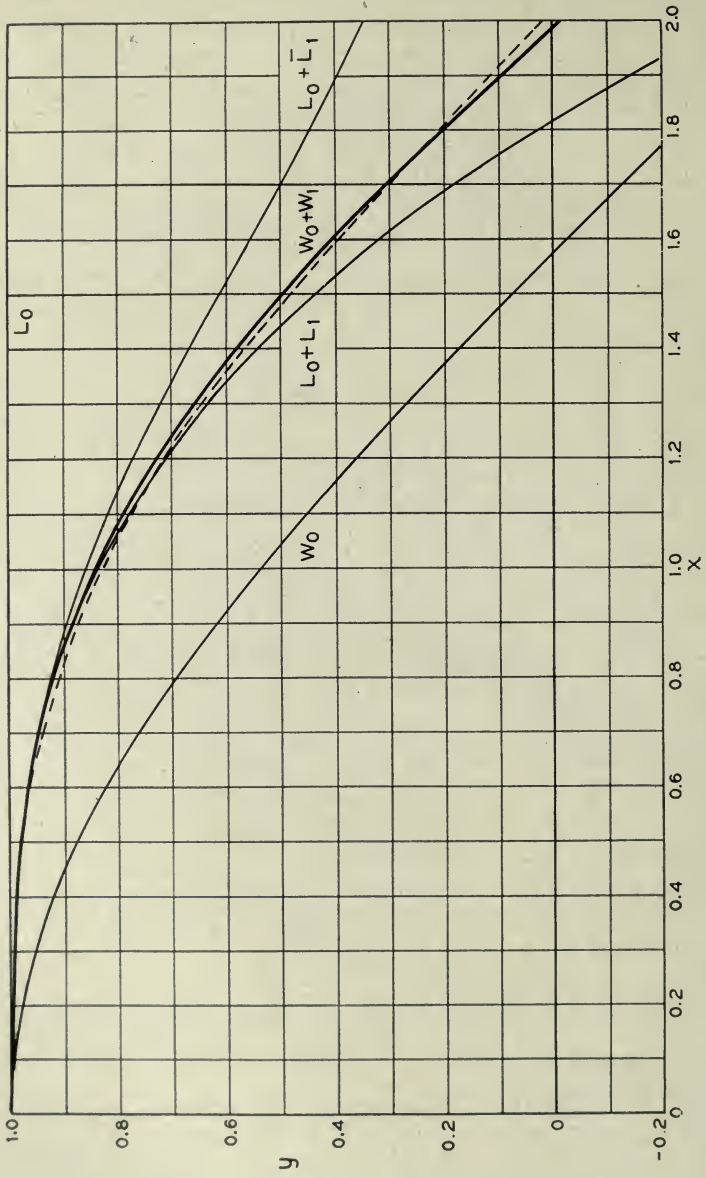


Fig. 1

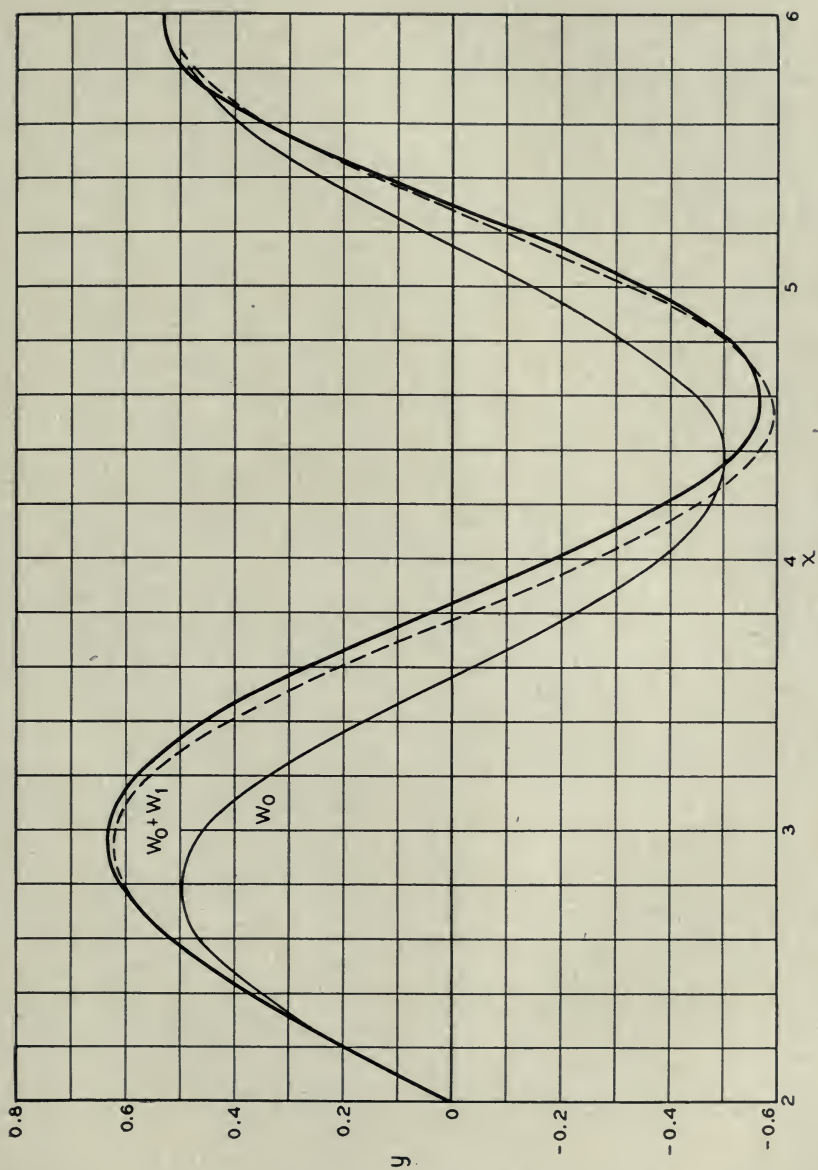


Fig. 2

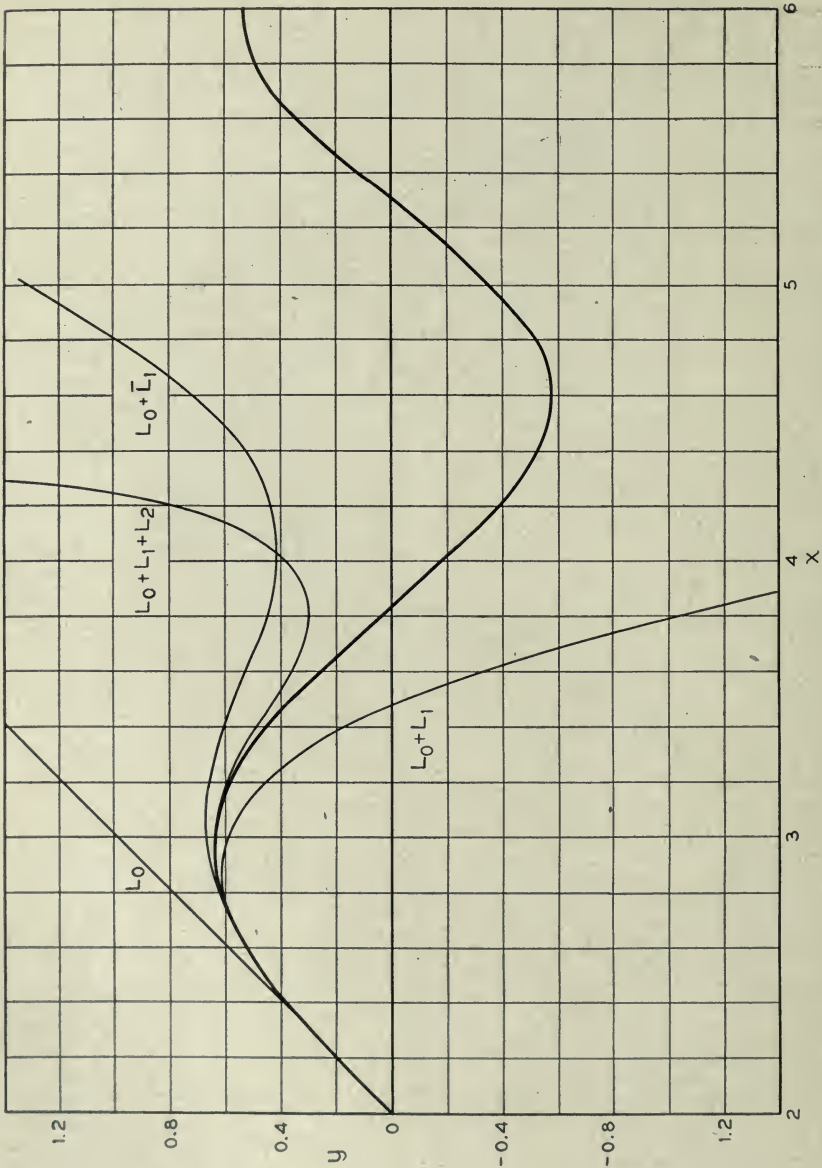


Fig. 3

(a) Wave perturbation, Fig. 2

$$W_0 = \frac{1}{2} \sin 2(x - 2)$$

$$W_1 = \frac{6-x}{32} \sin 2(x-2) + \frac{12-8x+x^2}{16} \cos 2(x-2).$$

Figure 2 exhibits rapid pulling of the successive approximate waves toward the exact even though the interval has been chosen deliberately unfavorable to the straight wave perturbation method [see example (c) and Fig. 4 for the improved treatment].

(b) Linear perturbation, Fig. 3

$$L_0 = x - 2$$

$$L_1 = \frac{1}{12} (16 - 16x + 4x^3 - x^4)$$

$$L_2 = -\frac{16}{63} + \frac{16x}{45} - \frac{2x^3}{9} + \frac{x^4}{9} - \frac{x^6}{90} + \frac{x^7}{504}.$$

Using W_0 instead of L_0

$$\bar{L}_1 = \frac{7}{8} - \frac{x}{2} + \frac{x}{8} \sin 2(x-2) + \frac{1}{8} \cos 2(x-2).$$

(c) Preliminary transformation of variables, Fig. 4

Introduce $\theta = \frac{2}{3}x^{3/2}$, $y = \theta^{-1/6}v$

and the modified equation is

$$v'' = -\left(1 + \frac{5}{36\theta^2}\right)v \quad \frac{4\sqrt{2}}{3} \leq \theta \leq 4\sqrt{6}.$$

Then, using for simplicity $\beta = 1$

$$v_0 = 2^{-1/2}\theta_0^{1/6} \sin(\theta - \theta_0), \theta_0 = 4\sqrt{2}/?$$

or

$$W_0 = (2x)^{-1/4} \sin \frac{2}{3}(x^{3/2} - 2^{3/2})$$

It will be seen that W_0 is a very good approximation throughout the range (2, 6). Adding W_1 obtained from

$$v_1 = \frac{5\theta_0^{1/6}}{36\sqrt{2}} [\cos(\theta + \theta_0)(Si\,2\theta - Si\,2\theta_0) - \sin(\theta + \theta_0)(Ci\,2\theta - Ci\,2\theta_0)]$$

the accurate curve y is reproduced.

In Fig. 2 the third approximation could not be distinguished from the accurate curve though numerically the values are not identical. For purposes of comparison the table of numerical values (Table A) may be found interesting.

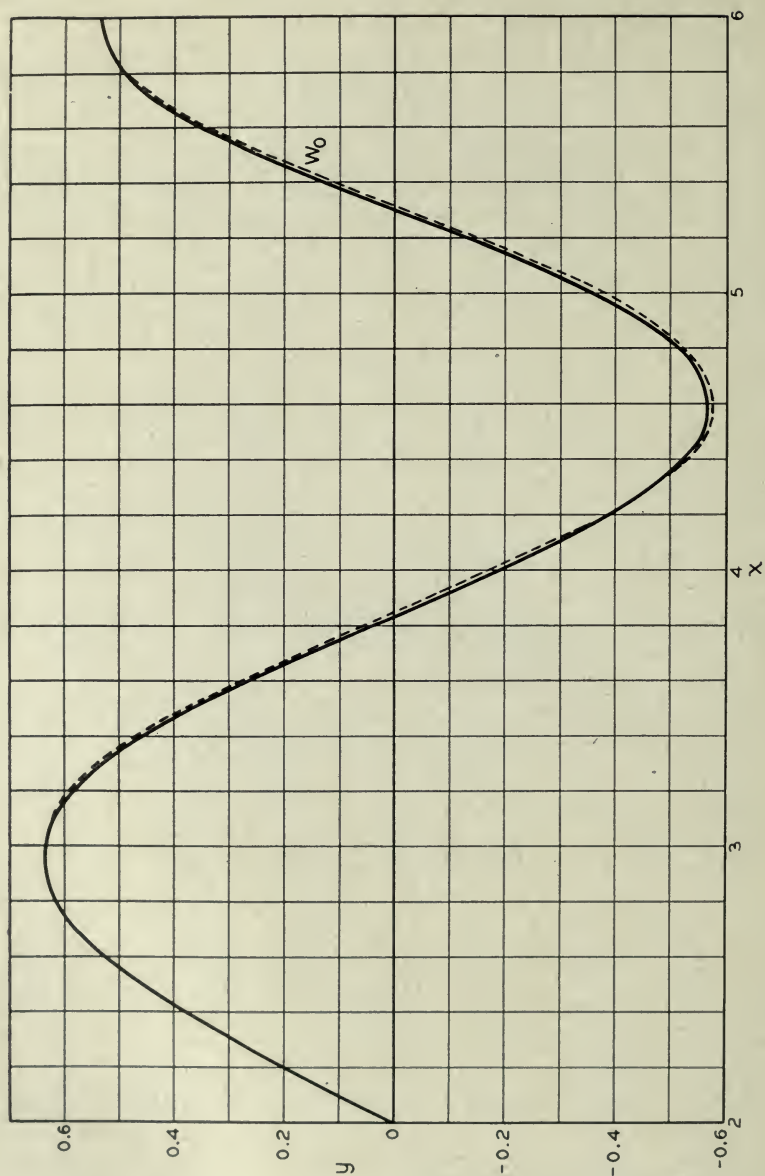


Fig. 4

TABLE A

x	y	W_0	$\frac{1}{0} \Sigma W$	$\frac{2}{0} \Sigma W$	$\frac{3}{0} \Sigma W$
2.0	0.	0.	0.	0.	0.
2.2	.19721	.19471	.19720	.19721	.19721
2.4	.37694	.35868	.37668	.37694	.37694
2.6	.52056	.46602	.51885	.52053	.52056
2.8	.61035	.49979	.60442	.61020	.61034
3.0	.63236	.45465	.61792	.63180	.63232
3.2	.57922	.33773	.55169	.57781	.57918
3.4	.45287	.16749	.40907	.44995	.45726
3.6	.26584	-.02919	.20603	.26085	.26561
3.8	.04126	-.22126	-.02974	.03408	.04087
4.0	-.18921	-.37840	-.26229	-.19800	-.18974
4.2	-.38951	-.47580	-.45326	-.39852	-.39011
4.4	-.52506	-.49808	-.56889	-.53240	-.52557
4.6	-.56943	-.44173	-.58697	-.57328	-.56964
4.8	-.51062	-.31563	-.50217	-.51000	-.51044
5.0	-.35548	-.13971	-.32847	-.35076	-.35494
5.2	-.13068	.05827	-.09772	-.12364	-.13006
5.4	-.12052	.24706	.14547	.12725	.12080
5.6	.34582	.39683	.35200	.34988	.34536
5.8	.49485	.48396	.47807	.49525	.49347
6.0	.53114	.49467	.49467	.52903	.52903

Example 3, Fig. 5

$$y'' = -y + \frac{2}{x^2}y, \quad 1 \leq x \leq \infty$$

$$y(1) = 1, y'(1) = 0$$

Exact solution: $y(x) = \sin(x-1) + \frac{1}{x} \cos(x-1)$

(a) Wave perturbation, with the initial conditions satisfied exactly

$$W_0 = \cos(x-1)$$

$$W_1 = 2 \sin(x-1) - 2 \cos(x+1) (Ci\ 2x - Ci\ 2) \\ - 2 \sin(x+1) (Si\ 2x - Si\ 2)$$

(b) Wave perturbation, matching the exact solution at infinity

$$\tilde{W}_0 = \sin(x-1)$$

$$\tilde{W}_1 = 2 \sin(x+1) Ci\ 2x - 2 \cos(x+1) (Si\ 2x - \pi/2)$$

This is an example of a solution in an infinite interval, where the perturbation term is not small throughout. It is interesting to note that the second form gives good agreement with the accurate solution in most of the range of integration.

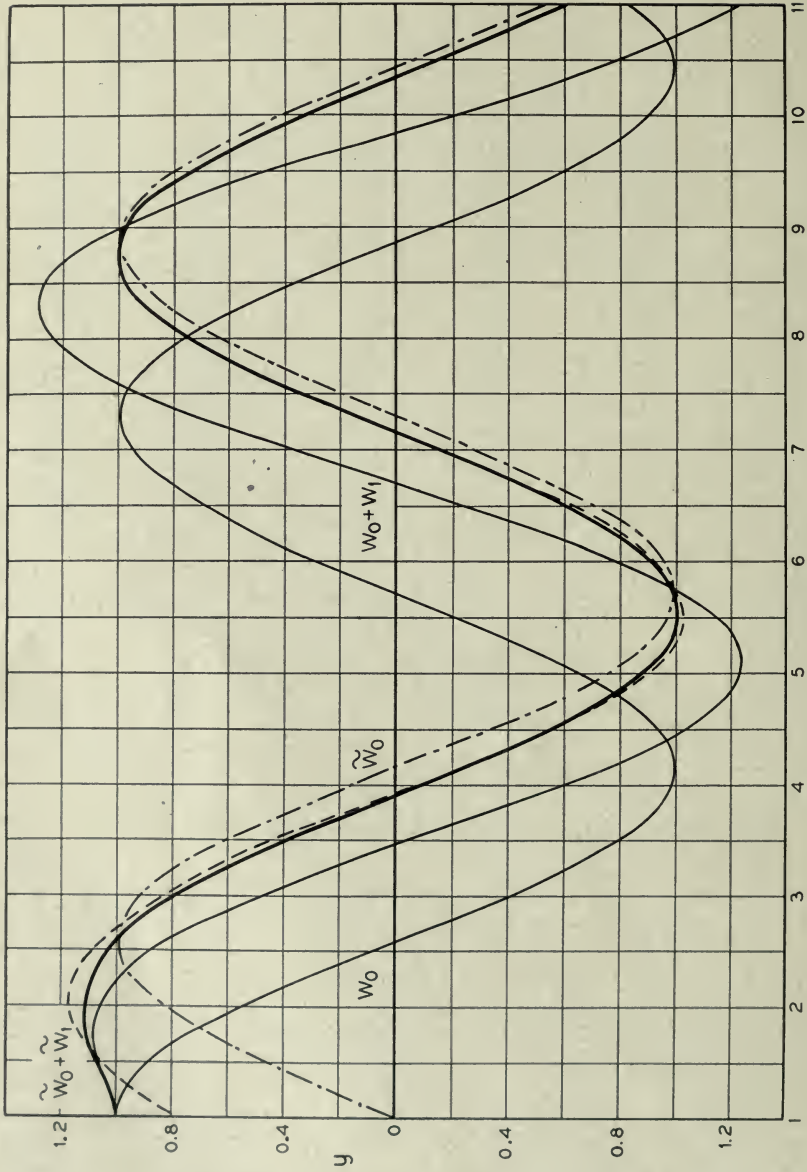


Fig. 5

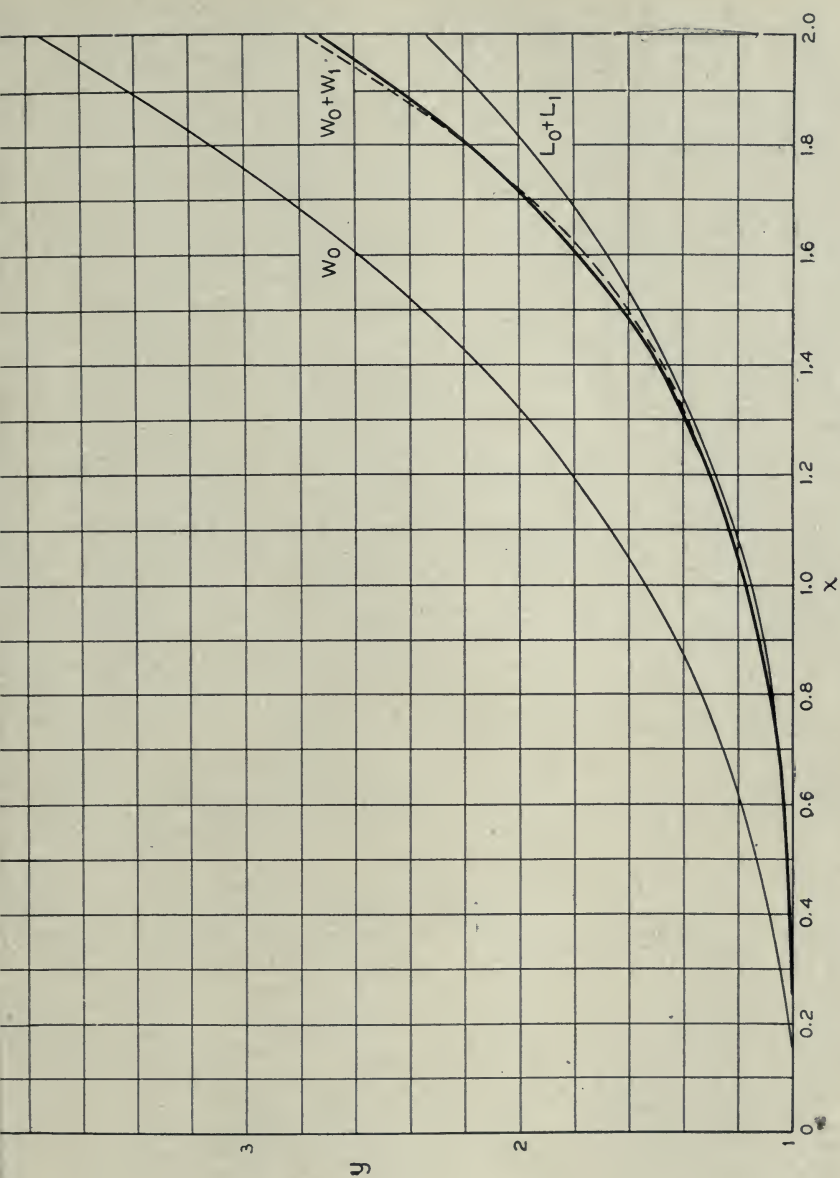


Fig. 6

Example 4, Fig. 6

$$y'' = +xy, \quad 0 \leq x \leq 2$$

$$y(0) = 1, y'(0) = 0$$

Exact solution:

$$y(x) = \Gamma(\frac{2}{3}) 3^{-1/3} x^{1/2} I_{-1/3}(\frac{2}{3} x^{3/2}).$$

(a) Wave perturbation

$$W_0 = \cosh x$$

$$W_1 = -\frac{x}{4} \cosh x + \frac{1}{4}(x^2 - 2x + 1) \sinh x$$

(b) Linear perturbation

$$L_0 = 1$$

$$L_1 = \frac{x^3}{6}.$$

This is an example in which the exact solution is non-oscillatory yet even in the short interval (0, 2) $W_0 + W_1$ is a better approximation than $L_0 + L_1$.

Example 5, Table I

$$y'' + \frac{1}{x} y' + y = 0, \quad 0 < x \leq \infty$$

Solution required to match the accurate solution

$$y(x) = J_0(x) - i N_0(x)$$

at infinity:

$$\tilde{W}_0 = \frac{1+i}{\sqrt{\pi x}} e^{-ix}$$

$$\tilde{W}_1 = \frac{1+i}{4\sqrt{\pi x}} e^{ix} \left[Ci2x - i \left(Si2x - \frac{\pi}{2} \right) \right].$$

TABLE I

x	$J_0 - iN_0$	$\tilde{W}_0$	$\tilde{W}_0 + \tilde{W}_1$	$\tilde{W}_0 + \tilde{W}_1 + \tilde{W}_2$
10	-.2459 -i .0557	-.2468 -i .0526	-.2460 -i .0557	
9	-.0903 -i .2499	-.0938 -i .2489	-.0903 -i .2500	
8	.1717 -i .2235	.1683 -i .2264	.1717 -i .2235	
7	.3001 +i .0259	.3009 +i .0207	.3001 +i .0260	
6	.1506 +i .2882	.1568 +i .2855	.1507 +i .2883	
5	-.1776 +i .3085	-.1704 +i .3135	-.1777 +i .3086	
4	-.3975 +i .0169	-.3979 +i .0291	-.3973 +i .0169	
3	-.2601 -i .3769	-.2765 -i .3684	-.2601 -i .3772	
2	.2239 -i .5104	.1967 -i .5288	.2246 -i .5109	
1	.7652 -i .0883	.7796 -i .1699	.7683 -i .0860	.7651 -i .0882
0.8	.8463 +i .0868	.8920 -i .0130	.8499 +i .0916	.8461 +i .0868
0.6	.9120 +i .3086	1.0124 +i .1899	.9152 +i .3183	.9116 +i .3084

For values of x less than 1 $\bar{W}_2$ was evaluated numerically.
Example 6, Table II

$$y'' + \frac{1}{x} y' = \left(\frac{1}{x^2} - 1 \right) y, \quad 1 \leq x \leq 3.$$

$$y(1) = 1, \quad y'(1) = 0.$$

The solution of this equation using Picard's method and the integrgraph has been described by Thornton C. Fry.³ We compare his results with those obtained by the wave perturbation method. The equation is first reduced to normal form by the substitution $y = x^{-1/2} u$, so that

$$u'' = \left(-1 + \frac{3}{4x^2} \right) u$$

and we have $\beta = \frac{1}{2}\sqrt{3}$. Then

$$x^{1/2} W_0 = \cos \beta (x - 1) + \frac{1}{2\beta} \sin \beta (x - 1)$$

$$x^{1/2} W_1 = \frac{1}{4\beta} \int_1^x \left(\frac{3}{u^2} - 1 \right) \left[\cos \beta (u - 1) + \frac{1}{2\beta} \sin \beta (u - 1) \right] \sin \beta (x - u) du.$$

While W_1 may be evaluated in terms of Ci and Si functions the values tabulated below were obtained by numerical integration. The values of the accurate solution

$$y = 1.4034 J_1(x) - 0.3251 N_1(x),$$

and of the third and eighth Picard approximations, are copied from Fry's paper.

TABLE II

x	y	y_3	y_8	W_0	$W_0 + W_1$
1.0	1.000	1.000	1.000	1.000	1.000
1.2	.998	.998	.998	.990	.998
1.4	.985	.984	.986	.961	.985
1.6	.956	.951	.955	.913	.956
1.8	.908	.894	.910	.848	.908
2.0	.842	.809	.844	.769	.842
2.2	.759	.694	.760	.677	.758
2.4	.659	.548	.661	.575	.659
2.6	.547	.370	.549	.466	.547
2.8	.425	.156	.427	.352	.427
3.0	.297	-.096	.300	.236	.300

REFERENCES

1. S. A. Schelkunoff, Solution of linear and slightly nonlinear equations, *Quart. App. Math.*, vol. 3, p. 348, Jan., 1946.
2. J. Liouville, Mémoires sur le développement des fonctions ou parties de fonctions en séries dont les divers termes sont assujétis à satisfaire à une même équation différentielle du second ordre, contenant un paramètre variable, *Jl. de Math: Pures et Appl.*, v. 1, p. 253-265, 1836, v. 2, p. 16-35 and p. 418-436, 1837. A brief discussion of the method will be found in "Numerical studies in differential equations" by H. Levy and E. A. Baggott, London, 1934; but these authors apply it to the numerical solution of non-linear equations and do not seem to have appreciated that its real potentialities lie in the field of linear equations, where they use only the better known methods.
3. Thornton C. Fry, The use of the integrator in the practical solution of differential equations by Picard's method of successive approximations, *Proc. Int. Math. Congress of Toronto*, pp. 405-428, 1924.

Potential Coefficients for Ground Return Circuits

By W. HOWARD WISE

This paper is concerned with the effect of the finite conductivity and dielectric constant of the earth on the potential coefficient for a 1-wire ground return circuit. It has been customary to say that the potential coefficient V/Q is

$$p_{12} = c^2 \log \rho''/\rho', \text{ elm units per cm.}$$

It is generally realized of course that this is just a good approximation to the true p_{12} . To see that it is just an approximation one has only to imagine the earth turning into air, in which case the distance ρ'' will eventually cease to have significance. The object of this paper is to derive the complete expression for p_{12} .

It turns out that

$$p_{12} = c^2 [2 \log \rho''/\rho' + 4(M + iN)] \quad (7)$$

where

$$M + iN = \int_0^\infty \frac{e^{-(h+z)\xi} \sqrt{\alpha} t \cos y \xi \sqrt{\alpha} t}{\sqrt{t^2 + i\epsilon i^2 \eta} + (\epsilon - i2c\lambda\sigma)t} dt \quad (8)$$

$\alpha = 4\pi\sigma\omega$, as in Carson's work on Z_{12} and in mine on Z_{12} at high frequencies^{1,2}

$$\xi e^{i\eta} = \sqrt{1 + i(\epsilon - 1)/2c\lambda\sigma} = s$$

ϵ = dielectric constant in electrostatic units

σ = conductivity in electromagnetic units

= 10^{-13} to 10^{-14} in ordinary soil

λ = wavelength in centimeters

c = velocity of light, in cm per sec.

$M + iN$ vanishes as $f \rightarrow 0, f \rightarrow \infty, \epsilon \rightarrow \infty$ or $\sigma \rightarrow \infty$.

Ordinarily $4(M + iN)$ will not be an important correction to $2 \log \rho''/\rho'$; but if the frequency is high and h or z is small it can be a worthwhile correction. For example, if a .02535 inch wire be thrown out on the ground to be a 2 mc antenna and we assume that $\sigma = 10^{-13}$, $\epsilon = 15$ and $h = 3$ cm. then, with a for wire radius,

$$\begin{aligned} p_{11} &= c^2 [2 \log 2h/a + 4(M + iN)] \\ &= c^2 [10.455 + .152 + i .319]. \end{aligned}$$

$1/p_{11}$ is the capacity to ground. If there were two parallel wires the scalar potential at the first wire would be $V_1 = p_{11}Q_1 + p_{12}Q_2$.

DERIVATION OF THE FORMULA

WE BEGIN with the wave-function for an exponentially propagated current in a straight wire parallel to a flat earth. The wave-function for a horizontal current-element dipole has been formulated as an infinite

¹ John R. Carson: "Wave Propagation in Overhead Wires with Ground Return," *Bell Sys. Tech. Jour.* 5, pp. 539-554, 1926.

² W. Howard Wise: "Propagation of High-Frequency Currents in Ground Return Circuits," *Proc. I. R. E.* 22, pp. 522-527, 1934.

integral by H. von Hoerschelmann³. The wave-function for the current in the wire is obtained by integrating the wave-functions of the current-element dipoles along the wire from minus infinity to plus infinity. It is

$$\begin{aligned} \Pi = & a \times e^{-\gamma x} \int_{-\infty}^{\infty} I_0 e^{-\gamma x} \left(\frac{e^{-ikR_1}}{R_1} - \frac{e^{-ikR_2}}{R_2} \right. \\ & \left. + \int_0^{\infty} \frac{2J_0(\nu\rho)}{l+m} e^{-wl} \nu \cdot d\nu \right) dx + b \times 0 \quad (1) \\ & - c \times 2e^{-\gamma x} \int_{-\infty}^{\infty} I_0 e^{-\gamma x} \frac{\partial}{\partial x} \int_0^{\infty} \frac{(1-\tau^2)J_0(\nu\rho)\nu}{(l+m)(l+\tau^2 m)} e^{-wl} d\nu \cdot dx. \end{aligned}$$

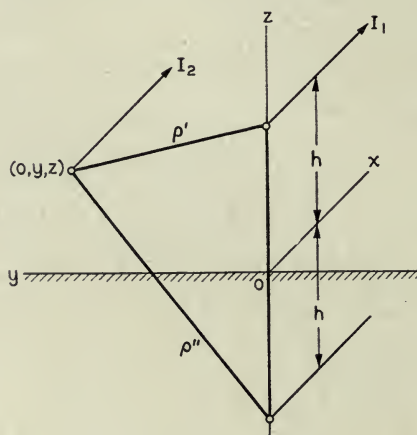


Fig. 1

The time factor is $e^{i\omega t}$. a , b and c are unit vectors pointing in the x , y and z directions.

$$\begin{aligned} R_1 &= (x^2 + y^2 + (h - z)^2)^{1/2} \\ R_2 &= (x^2 + y^2 + (h + z)^2)^{1/2} \\ \rho &= (x^2 + y^2)^{1/2} \\ k &= 2\pi/\lambda \\ k_2^2 &= \epsilon\mu\omega^2 - i4\pi\sigma\mu\omega \text{ in electromagnetic units} \end{aligned}$$

By supposing ϵ to be measured in electrostatic units, we can write

$$k_2^2 = k^2(\epsilon - i2c\lambda\sigma)\mu.$$

³ H. Von Hoerschelmann, Jahrb. der draht. Teleg. 5, pp. 14-188, 1912.

It is assumed that μ is everywhere unity in electromagnetic units.

$$\begin{aligned} l &= (\nu^2 - k^2)^{1/2}, & m &= (\nu^2 - k_2^2)^{1/2} \\ w &= h + z, & \tau^2 &= k^2/k_2^2 \\ \gamma &= \alpha + i\beta \text{ is the desired propagation constant} \end{aligned}$$

The electric field parallel to the wire is

$$\begin{aligned} E_x &= -i\omega \left[\Pi_x + k^{-2} \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} \Pi_x + \frac{\partial}{\partial y} \Pi_y + \frac{\partial}{\partial z} \Pi_z \right) \right] \\ &= -I_0 e^{-\gamma x} Z_{12} - \frac{\partial V}{\partial x} \end{aligned} \quad (2)$$

It has previously been shown that²

$$Z_{12} = i\omega [2 \log \rho''/\rho' + 4(Q - iP)] \quad (3)$$

where

$$\begin{aligned} Q - iP &= \frac{1}{is^2} \int_0^\infty (\sqrt{\nu^2 + is^2} - \nu) e^{-w'\nu} \cos y'\nu \cdot d\nu, \\ w' &= w \sqrt{\alpha} \text{ and } y' = y \sqrt{\alpha}. \end{aligned}$$

To get the potential coefficient for a ground return circuit it is necessary to compute the scalar potential.

$$\begin{aligned} V &= i\omega k^{-2} \left(\frac{\partial}{\partial x} \Pi_x + \frac{\partial}{\partial y} \Pi_y + \frac{\partial}{\partial z} \Pi_z \right) \\ &= Q p_{12} \end{aligned} \quad (4)$$

As in previous work the propagation constant γ is assigned the value ik as a first approximation. This is an ideal value for γ but the following considerations make it an imperative choice: (1) to assume that the current is propagated down the line with a velocity less than that of light makes the integrals very hard to evaluate, (2) to assume that the attenuation is not zero on an infinite line amounts to assuming an infinite source of energy and makes the integrals diverge.

It should not be inferred that the resulting formulas are necessarily poor if the physical system does not closely approximate the ideal one in which γ is ik . ik is employed as a convenient first approximation in evaluating the correction terms in Z_{12} and p_{12} . Eventually, if there were but one wire, one would compute $\gamma = \sqrt{(z + Z_{11})(G + i\omega/p_{11})}$, wherein Z_{11} and p_{11} have been evaluated with ik for γ , and this would be a second approximation to γ . Past experience with the second approximation so obtained has justified the expectation that it would be a satisfactory final result. Since

the integrals diverge if the attenuation is not zero the use of an infinite line formula presupposes reasonably efficient transmission.

Since $\Pi \propto e^{-\gamma x}$ we have

$$i\omega k^{-2} \frac{\partial}{\partial x} \Pi_x = \frac{\omega}{k} \Pi_x = \frac{I}{ik} Z_{12} = \frac{cI}{i\omega} Z_{12}.$$

Since $-\frac{\partial I}{\partial x} = \frac{\partial Q}{\partial t}$ or $ikI = i\omega Q$ or $I = cQ$ this is

$$i\omega k^{-2} \frac{\partial}{\partial x} \Pi_x = Qc^2 [2 \log \rho''/\rho' + 4(Q - iP)]. \quad (5)$$

We have next to consider

$$\begin{aligned} \frac{i\omega}{k^2} \frac{\partial}{\partial z} \Pi_z &= \frac{2\omega(1 - \tau^2)}{ik^2} I \frac{\partial}{\partial z} \int_{-\infty}^{\infty} e^{-\gamma x} \frac{\partial}{\partial x} \int_0^{\infty} \frac{J_0(\nu\rho) e^{-w l} \nu \cdot d\nu}{(l+m)(l+\tau^2 m)} dx \\ &= Qc^2 \frac{2(1 - \tau^2)}{ik} \frac{\partial}{\partial z} \int_0^{\infty} \frac{e^{-w l} \nu \cdot d\nu}{(l+m)(l+\tau^2 m)} \int_{-\infty}^{\infty} e^{-\gamma x} \frac{\partial}{\partial x} J_0(\nu\rho) dx. \end{aligned}$$

The infinite integral is

$$\int_0^{\infty} \frac{e^{-w l} \nu \cdot d\nu}{(l+m)(l+\tau^2 m)} \left[e^{-\gamma x} J_0(\nu\rho) \right]_{-\infty}^{\infty} + \gamma \int_{-\infty}^{\infty} e^{-\gamma x} J_0(\nu\rho) dx \Bigg].$$

Since $J_0(\nu \sqrt{x^2 + y^2})$ and $\cos kx$ are even functions of x and $\sin kx$ is an odd function of x

$$\begin{aligned} \int_{-\infty}^{\infty} J_0(\nu \sqrt{x^2 + y^2}) e^{-ikx} dx &= 2 \int_0^{\infty} J_0(\nu \sqrt{x^2 + y^2}) \cos kx \cdot dx \\ &= 0 \quad \text{if } \nu < k \\ &= 2 \frac{\cos y \sqrt{\nu^2 - k^2}}{\sqrt{\nu^2 - k^2}} \quad \text{if } k \leq \nu \end{aligned}$$

and so our integral is

$$2ik \int_k^{\infty} \frac{e^{-w l} \nu}{(l+m)(l+\tau^2 m)} \cdot \frac{\cos y l}{l} d\nu$$

or, since $l^2 = \nu^2 - k^2$,

$$2ik \int_0^{\infty} \frac{e^{-w l} \cos y l \cdot dl}{(l + \sqrt{l^2 + i^2(k_2^2 - k^2)})(l + \tau^2 \sqrt{l^2 + i^2(k_2^2 - k^2)})}$$

or, if we put $l = \nu \sqrt{\alpha}$

$$\text{and } i(k_2^2 - k^2) = 4\pi\sigma\omega \left(1 + i \frac{\epsilon - 1}{2c\lambda\sigma}\right) = \alpha s^2,$$

$$\frac{2k}{\sqrt{\alpha}} \int_0^\infty \frac{e^{-w'\nu} \cos y'\nu \cdot d\nu}{(\nu + \sqrt{\nu^2 + is^2})(\nu + \tau^2 \sqrt{\nu^2 + is^2})}$$

where, as in $Q - iP$, $w' = w \sqrt{\alpha}$ and $y' = y \sqrt{\alpha}$.

Noting next that $\frac{\partial}{\partial z} e^{-w'\nu} = -\sqrt{\alpha\nu} e^{-w'\nu}$

we have

$$\begin{aligned} \frac{i\omega}{k^2} \frac{\partial}{\partial z} \Pi_z &= -Qc^2 4 \int_0^\infty \frac{(1 - \tau^2)e^{-w'\nu} \cos y'\nu \cdot d\nu}{(\nu + \sqrt{\nu^2 + is^2})(\nu + \tau^2 \sqrt{\nu^2 + is^2})} \\ &= -Qc^2 \frac{4}{is^2} \int_0^\infty \frac{\sqrt{\nu^2 + is^2} - \nu}{\nu + \tau^2 \sqrt{\nu^2 + is^2}} e^{-w'\nu} \cos y'\nu \cdot (1 - \tau^2)\nu \cdot d\nu. \end{aligned}$$

Since $(1 - \tau^2)\nu = \nu + \tau^2 \sqrt{\nu^2 + is^2} - \tau^2(\sqrt{\nu^2 + is^2} + \nu)$

this is

$$\begin{aligned} -Qc^2 \frac{4}{is^2} \int_0^\infty \left[\sqrt{\nu^2 + is^2} - \nu - \frac{\tau^2 is^2}{\nu + \tau^2 \sqrt{\nu^2 + is^2}} \right] e^{-w'\nu} \cos y'\nu \cdot d\nu \\ = Qc^2 \left[-4(Q - iP) + 4 \int_0^\infty \frac{e^{-w'\nu} \cos y'\nu \cdot d\nu}{\sqrt{\nu^2 + is^2} + \nu/\tau^2} \right]. \quad (6) \end{aligned}$$

On adding (6) to (5) we have

$$Qp_{12} = Qc^2[2 \log \rho''/\rho' + 4(M + iN)] \quad (7)$$

where

$$\begin{aligned} M + iN &= \int_0^\infty \frac{e^{-w'\nu} \cos y'\nu}{\sqrt{\nu^2 + is^2} + \nu/\tau^2} d\nu \\ &= \int_0^\infty \frac{e^{-(h+z)\zeta} \sqrt{\alpha} t \cos y \sqrt{\alpha} \zeta t}{\sqrt{t^2 + ie^{i2\eta}} + t/\tau^2} dt. \quad (8) \end{aligned}$$

$M + iN$ vanishes as $f \rightarrow 0$, $f \rightarrow \infty$, $\epsilon \rightarrow \infty$ or $\sigma \rightarrow \infty$.

When $k_2^2 - k^2$ is minute the leading terms in the approximation (9)

for $M + iN$ are $\frac{\pi}{8} - \frac{C}{2} - \frac{1}{2} \log(p'' \sqrt{(k_2^2 - k^2)/2}) - i \frac{\pi}{4}$.

AN APPROXIMATION FOR $M + iN$

It is possible to get series expansions for $M + iN$ but those which have been obtained do not facilitate computation. A fairly good approximation to $M + iN$ is arrived at as follows.

$$\text{Let } ie^{i2\eta} = u^2, \quad u = e^{i(\eta+\pi/4)}$$

$$\epsilon - i2c\lambda\sigma = a = 1/\tau^2$$

$$e^{-(h+z)\zeta\sqrt{\alpha}t} \cos y\zeta\sqrt{\alpha}t = \frac{1}{2}(e^{-a't} + e^{-a''t})$$

$$g' = (h+z-iy)\zeta\sqrt{\alpha}$$

$$g'' = (h+z+iy)\zeta\sqrt{\alpha}$$

$$g = (h+z)\zeta\sqrt{\alpha}$$

$$\begin{aligned} \text{Since } (t^2 + u^2)^{1/2} &= (t^2 + 2tu + u^2 - 2tu)^{1/2} \\ &= t + u - tu/(t+u) + \dots \end{aligned}$$

we put

$$1/(\sqrt{t^2 + u^2} + at) = 1/(t+u - tu/(t+u) + at)$$

$$= (t+u)/[(a+1)t^2 + (a+1)tu + u^2]$$

$$= (t+r_1+r_2)/(a+1)(t+r_1)(t+r_2)$$

where $r_1 = u(1 - \sqrt{1 - 4/(a+1)})/2$.

and $r_2 = u(1 + \sqrt{1 - 4/(a+1)})/2$.

Then

$$\begin{aligned} M + iN &\approx \frac{1}{(a+1)(r_2 - r_1)} \int_0^\infty \left(\frac{r_2}{t+r_1} - \frac{r_1}{t+r_2} \right) \frac{e^{-a't} + e^{-a''t}}{2} dt \\ &= \frac{1}{2(a+1)(r_2 - r_1)} [-r_2 e^{a'r_1} \text{li}(e^{-a'r_1}) + r_1 e^{-a'r_2} \text{li}(e^{-a'r_2}) \\ &\quad - r_2 e^{a''r_1} \text{li}(e^{-a''r_1}) + r_1 e^{a''r_2} \text{li}(e^{-a''r_2})] \quad (9) \end{aligned}$$

When $y = 0$ this reduces to

$$M + iN \approx \frac{1}{(a+1)(r_2 - r_1)} [-r_2 e^{a'r_1} \text{li}(e^{-a'r_1}) + r_1 e^{a'r_2} \text{li}(e^{-a'r_2})]. \quad (10)$$

$$\text{li}(e^{-z}) = - \int_z^\infty \frac{e^{-t}}{t} dt = C + \log z + \sum_{i=1}^\infty \frac{(-z)^i}{i!t},$$

where $C = .577215665$ and, if $z = re^{i\theta}$, $-\pi \leq \theta \leq \pi$.

$$\text{li}(e^{-z}) \sim \frac{e^{-z}}{-z} \sum_{i=0}^n \frac{i!}{(-z)^i} + R_n.$$

The accompanying charts give M and N for $y = 0$ and $\epsilon = 15$. With $z = h$ they give M and N for p_{11} . The computed points are indicated by solid dots on the chart for M ; they were obtained by numerical integration.

The approximation (10) was checked against the values obtained by numerical integration at a number of points. The discrepancy in each case amounted to less than one per cent for both M and N . This approximation is a much easier way to evaluate the integral than is numerical integration but it is a tedious computation with many chances for error. Conse-

quently it is important to observe that a coarser kind of approximation may often be good enough. Thus, taking y to be zero,

$$M + iN \approx \int_0^\infty \frac{e^{-gt}}{u + at} dt = -\frac{1}{a} e^{gu/a} \operatorname{li}(e^{-gu/a}).$$

If g is very small one might use

$$\begin{aligned} M + iN &\approx \int_0^1 \frac{dt}{u + at} + \int_1^\infty \frac{e^{-gt}}{(1+a)t} dt \\ &= \frac{1}{a} \log \left(1 + \frac{a}{u} \right) - \frac{1}{1+a} \operatorname{li}(e^{-g}). \end{aligned}$$

Ordinarily precision is not required in $M + iN$ because $4(M + iN)$ is a small term in p_{12} .

Abstracts of Technical Articles by Bell System Authors

*Electrochemical Factors in Underground Corrosion of Lead Cable Sheath.*¹

V. J. ALBANO. Stray current is the principle cause of corrosion failures on underground telephone cables in most cities where trolleys are operated. To mitigate this condition, the cable sheaths are "drained" to the negative return system of the traction system. By this means not only is the stray current anodic area largely eliminated, but the cables automatically become negative to earth, and therefore are cathodically protected. The cathodic protection afforded in this manner prevents other types of corrosion from occurring. With the gradual abandonment of trolley systems, and with the extension of underground cables into non-trolley areas, the percentage of underground telephone plant receiving this protection is decreasing. As a result, the problems of lead corrosion due to such causes as galvanic and local cell action of various types, and chemical action by substances in the soil are becoming more prevalent. It is the purpose of this article to review some of the basic principles of corrosion not involving stray currents, and show how they apply to the problems of lead cable sheath corrosion.

*PCM Equipment.*² H. S. BLACK and J. O. EDSON. PCM, pulse code modulation, is a new solution to the problem of overcrowded frequency spectrum. It appears to have exceptional possibilities from the standpoint of freedom from interference, and seems to have inherent advantages over other types of multiplexing.

*Coaxial-Cable Networks.*³ FRANK A. COWAN. This paper discusses the general features of the coaxial system, its application for both telephone and television, and the future prospects for very-broadband transmission facilities in the communication network.

*Parabolic-Antenna Design for Microwaves.*⁴ C. C. CUTLER. This paper is intended to give fundamental relations and design criteria for parabolic radiators at microwave frequencies (i.e., wavelengths between 1 and 10 centimeters). The first part of the paper discusses the properties of the parabola which make it useful as a directional antenna, and the relation of phase polarization and amplitude of primary illumination to the over-all radiation characteristics. In the second part, the characteristics of practical feed systems for parabolic antennas are discussed.

¹ *Corrosion*, October 1947.

² *Electrical Engineering*, November 1947.

³ *Proc. I. R. E.—Waves and Electrons Section*, November 1947.

⁴ *Proc. I. R. E.*, November 1947.

*Microwave Antenna Measurements.*⁵ C. C. CUTLER, A. P. KING and W. E. KOCK. A description is given of the techniques involved in measuring the properties of microwave antennas. The measuring methods which are peculiar to these frequencies are discussed, and include the measurement of gain, beam width, minor lobes, wide-angle radiation, mutual coupling between antennas, phase, and polarization. The requirements of the antenna testing site are taken up, and components of a complete measuring system are briefly described.

*Microwave Converters.*⁶ C. F. EDWARDS. Microwave converters using point-contact silicon rectifiers as the nonlinear element are discussed, with particular emphasis on the design of the networks connecting the rectifier to the input and output terminals. Several converters which have been developed during recent years for use at wavelengths between 1 and 30 centimeters are described, and some of the effects of the impedance-versus-frequency characteristics of the networks on the converter performance are discussed.

*Recent Developments in Relays:*⁷ *Glass-Enclosed Reed Relay*, W. B. ELLWOOD; *Mercury Contact Relays*, J. T. L. BROWN and C. E. POLLARD. Relays which combine high-speed and great uniformity of performance over long periods of time are required for some uses in the telephone plant. The relays described possess these qualities to an unusual degree. Detailed description is limited to two types, each typical of a generic family in which the principles involved apply to all.

These relays are based on the philosophy that a motor element (any device for conversion of electromagnetic to mechanical energy), which is efficient and magnetically and elastically stable and operates contacts sealed in a proper atmosphere free from dirt and film, will give reliable performance if the contact load is engineered to the capacity of the contact. The relays require no maintenance beyond unit replacement, for there is no possibility of a change in adjustment after assembly is completed.

In one form the contact is provided for by metal in solid form, while in the other a mercury film supported on solid metal surfaces provides the contacting medium. The mercury at the contacting surfaces is replenished continuously through a capillary path from a mercury reservoir below the contact.

*An Adjustable Wave-Guide Phase Changer.*⁸ A. GARDNER FOX. A very interesting and useful component of the wave-guide art is the differential phase-shift section, wherein dominant waves of one polarization are caused to travel through a section of wave guide at a different velocity than waves

⁵ *Proc. I. R. E.*, December 1947.

⁶ *Proc. I. R. E.*, November 1947.

⁷ *Elec. Engg.*, November 1947.

⁸ *Proc. I. R. E.*, December 1947.

polarized at right angles to the first. Particularly useful are the $\Delta 90$ -degree and $\Delta 180$ -degree differential phase-shift sections which produce differential delays between the two polarizations of 90 degrees and 180 degrees, respectively. The properties of these sections are discussed, and it is shown how they may be combined to form a phase changer which will transmit substantially 100 per cent of the incident power with a phase which is readily adjustable. Several different methods of building these sections are finally described.

*Considerations in the Design of a Radar Intermediate-Frequency Amplifier.*⁹ ANDREW L. HOPPER and STEWART E. MILLER. The intermediate-frequency amplifier of a microwave radar receiver is commonly required to provide approximately 100 decibels amplification in a bandwidth of 1 to 10 megacycles, centered at frequencies in the 30- and 60-megacycle regions. Meeting such requirements involves the use of five to ten amplifier stages of the highest efficiency that can be suited to production methods. In addition, the noise figure of the radar intermediate-frequency amplifier is a significant contributor to the over-all radar receiver noise figure, and must therefore be maintained at an absolute minimum. By examining a particular intermediate-frequency-amplifier design (one providing an over-all bandwidth of 10 megacycles centered at 60 or 100 megacycles), this paper discusses qualitatively the theoretical problems involved in such a design and gives data of practical importance to the engineer attempting to build a similar amplifier. Measured characteristics of approximately fifty amplifiers are summarized to illustrate the end results achieved.

*Historical Note on the Rate of a Moving Atomic Clock.*¹⁰ HERBERT E. IVES. The history of the idea of variation of frequency with velocity is followed through Goigt, Larmor, Lorentz, and Einstein. The Michelson-Morley experiment is explainable by any contraction of dimensions in the ratio $(1 - v^2/c^2)^{1/2}:1$ along and transverse to the direction of motion. To each contraction corresponds a different value of frequency change. The theoretical speculations pointing to the relation $v_m = v_0(1 - v^2/c^2)^{1/2}$ are discussed, together with the significance of the experimental test by means of canal rays.

*New Low-Coefficient Synthetic Piezoelectric Crystals for Use in Filters and Oscillators.*¹¹ W. P. MASON. Two crystals of the monoclinic sphenoidal class have been found which have modes of vibration with zero temperature coefficients of frequency, high electromechanical coupling constants, and high Q 's or low dissipation. These properties make it appear probable that such crystals may have a considerable use in filters and oscillators as a sub-

⁹ *Proc. I. R. E.*, November 1947.

¹⁰ *Jour. Opt. Soc. Amer.*, October 1947.

¹¹ *Proc. I. R. E.*, October 1947.

stitute for quartz, which is difficult to obtain in large sizes. These crystals are ethylene diamine tartrate (EDT) having the chemical formula $C_6H_{14}N_2O_6$, and di-potassium tartrate (DKT) having the formula $K_2C_4H_4O_6 - \frac{1}{2} H_2O$.

The paper describes the properties of EDT, since this crystal has been found more advantageous than DKT. The 13 elastic constants, the 8 piezoelectric constants, and the 4 dielectric constants have been measured over a temperature range, and from these measurements the regions of low temperature coefficients and high electromechanical coupling have been located. Six low-temperature-coefficient cuts have been discovered and the properties of these cuts are given. These cuts are being applied in the crystal channel filters of the long-distance telephone system, and may be applied to the control of oscillators.

*Multi-Channel Carrier Telegraph.*¹² A. L. MATTE. Discussion of a carrier telegraph system, adapted specifically to railway requirements, to meet the needs for high-quality line transmission.

*Reflex Oscillators for Radar System.*¹³ J. O. McNALLY and W. G. SHEPHERD. The advantages to be gained in the operation of radar systems at very high frequencies have led to the use of frequencies of several thousand megacycles. Operation at these frequencies has imposed serious problems in obtaining suitable tube behavior. Because of the difficulty in obtaining amplification at the transmission frequency, the r.f. section of the usual radar receiver consists of a crystal converter driven by a beating oscillator and operating directly into an i.f. amplifier. Since the midband frequency of the latter has commonly been either 30 or 60 Mc., it has been necessary to provide beating oscillators operating at frequencies differing from those of the transmitter by only a few per cent.

For radar systems intended to operate at approximately 3000 Mc., which were under development in the early days of the war, it was found that triodes then available gave unsatisfactory performance. Attention shifted to the possibility of using velocity-modulated tubes, and the particular form known as the reflex oscillator came into general use.

In this paper the requirements on beating-oscillator tubes for radar systems are discussed, and the design features which have made the reflex oscillator eminently satisfactory in this application are pointed out. Problems encountered in such oscillators are outlined, and the solution in a number of cases is indicated. In some instances military requirements and expediency were in conflict with the optimum performance, and hence certain compromises were necessary.

¹² *Railway Signaling*, December 1947.

¹³ *Proc. I. R. E.*, December 1947.

*Space-Charge and Transit-Time Effects on Signal and Noise in Microwave Tetrodes.*¹⁴ L. C. PETERSON. Signal and noise in microwave tetrodes are discussed with particular emphasis on their behavior as space-charge conditions are varied in the grid-screen, or drift, region. The analysis assumes that the electron-stream velocity is single-valued. For particular conditions the noise figure may be substantially improved by increasing the space-charge density in the grid-screen region until an entering electron encounters a field of a certain magnitude. The noise reduction is largely due to the cancellation in the output of the noise produced by the random cathode emission. The method of noise reduction described is applicable only when the transit angles of both input and drift regions are fairly long.

In a forthcoming paper, H. V. Neher describes experimental results which broadly agree with the theory.

*"Cloverleaf" Antenna for F. M. Broadcasting.*¹⁵ PHILLIP H. SMITH. The radiation requirements and general design considerations for transmitting antennas suitable for f.m. broadcasting are briefly discussed, and an explanation of the design and operation of the arrangement of radiating elements and associated feed system employed in the "cloverleaf" antenna is given. Both calculated and measured data are included, showing field-intensity distribution, gain, impedance-frequency characteristics, etc. Design features which are discussed include a simple coaxial impedance-matching transformer developed initially for microwave application, and the method and facilities provided for the removal of sleet.

*Hybrid Circuits for Microwaves.*¹⁶ W. A. TYRRELL. The fundamental behavior of hybrid circuits is reviewed and discussed, largely in terms of reciprocity relationships. The phase properties of simple wave-guide tee junctions are briefly considered. Two kinds of hybrid circuits are then described, the one involving a ring or loop of transmission line, the other relying upon the symmetry properties of certain four-arm junctions. The description is centered about wave-guide structures for microwaves, but the principles may also be applied to other kinds of transmission lines for other frequency ranges. Experimental verification is provided, and some of the important applications are outlined.

¹⁴ *Proc. I. R. E.*, November 1947.

¹⁵ *Proc. I. R. E.—Waves and Electrons Section*, December 1947.

¹⁶ *Proc. I. R. E.*, November 1947.

Contributors to this Issue

J. R. DAVEY, B.S. in Electrical Engineering, University of Michigan, 1936. During the war Mr. Davey was engaged in the development of H. F. radio teletype systems, and has since been concerned with the development of electronic telegraph circuits.

H. T. FRIIS, E.E., Royal Technical College, Copenhagen, 1916; Sc.D., 1938; Assistant to Professor P. D. Pedersen, 1916; Technical Advisor at the Royal Gun Factory, Copenhagen, 1917-18; Fellow of the American Scandinavian Foundation, 1919; Columbia University, 1919. Western Electric Company, 1920-25; Bell Telephone Laboratories, 1925-. Formerly as Radio Research Engineer and since January 1946 as Director of Radio Research, Dr. Friis has long been engaged in work concerned with fundamental radio problems. He is a Fellow of the Institute of Radio Engineers.

MARION C. GRAY, Edinburgh University, M.A., 1922; Bryn Mawr College, Ph.D., 1926. Instructor in physics, Edinburgh University, 1926-27; Research Assistant in Mathematics, Imperial College, London, 1927-30. American Telephone and Telegraph Company, Department of Development and Research, 1930-34; Bell Telephone Laboratories, 1934-. Dr. Gray has been engaged mainly in mathematical work in the field of electromagnetic theory.

A. L. MATTE, B.S. in Electrical Engineering, Massachusetts Institute of Technology, 1909; Graduate Studies, M.I.T., 1912-13. New England Investment and Securities Company, 1910-12; Detroit United Railways, 1913-18. American Telephone and Telegraph Company, Department of Development and Research, 1918-34; Bell Telephone Laboratories, 1934-. Mr. Matte has been engaged principally in transmission studies relating to telegraphy.

S. O. RICE, B.S. in Electrical Engineering, Oregon State College, 1929; California Institute of Technology, 1929-30, 1934-35. Bell Telephone Laboratories, 1930-. Mr. Rice has been concerned with various theoretical investigations relating to telephone transmission theory.

D. H. RING, A.B., Stanford, 1929; Engineer, Stanford, 1930. Bell Telephone Laboratories, 1930-. Mr. Ring has been engaged in radio research.

S. A. SCHELKUNOFF, B.A., M.A. in Mathematics, The State College of Washington, 1923; Ph.D. in Mathematics, Columbia University, 1928. Engineering Department, Western Electric Company, 1923-25; Bell Telephone Laboratories, 1925-26. Department of Mathematics, State College of Washington, 1926-29. Bell Telephone Laboratories, 1929-. Dr. Schelkunoff has been engaged in mathematical research, especially in the field of electromagnetic theory.

W. H. WISE, B.S., Montana State College, 1921; M.A., University of Oregon, 1923; Ph.D., California Institute of Technology, 1926. American Telephone and Telegraph Company, Department of Development and Research, 1926-34; Bell Telephone Laboratories, 1934-. Dr. Wise has been engaged in various theoretical investigations relating to transmission theory.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS
OF ELECTRICAL COMMUNICATION

A Mathematical Theory of Communication. . . *C. E. Shannon* 379

An Aspect of the Dialing Behavior of Subscribers and Its
Effect on the Trunk Plant. *Charles Clos* 424

Spectra of Quantized Signals *W. R. Bennett* 446

Analysis and Performance of Waveguide-Hybrid Rings
for Microwaves. *H. T. Budenbom* 473

Methods of Electromagnetic Field Analysis
S. A. Schelkunoff 487

The Evolution of the Quartz Crystal Clock
Warren A. Marrison 510

Abstracts of Technical Articles by Bell System Authors. . 589

Contributors to this Issue 591

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

THE BELL SYSTEM TECHNICAL JOURNAL

*Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.*



EDITORS

R. W. King

J. O. Perrine

EDITORIAL BOARD

C. F. Craig

O. E. Buckley

O. B. Blackwell

M. J. Kelly

H. S. Osborne

A. B. Clark

J. J. Pilliod

F. J. Feely



SUBSCRIPTIONS

Subscriptions are accepted at \$1.50 per year. Single copies are 50 cents each.
The foreign postage is 35 cents per year or 9 cents per copy.



Copyright, 1948
American Telephone and Telegraph Company

The Bell System Technical Journal

Vol. XXVII

July, 1948

No. 3

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance

¹ Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A. I. E. E. Trans.*, v. 47, April 1928, p. 617.

² Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal*, July 1928, p. 535.

such as time, bandwidth, number of relays, etc., tend to vary linearly with the logarithm of the number of possibilities. For example, adding one relay to a group doubles the number of possible states of the relays. It adds 1 to the base 2 logarithm of this number. Doubling the time roughly squares the number of possible messages, or doubles the logarithm, etc.

2. It is nearer to our intuitive feeling as to the proper measure. This is closely related to (1) since we intuitively measure entities by linear comparison with common standards. One feels, for example, that two punched cards should have twice the capacity of one for information storage, and two identical channels twice the capacity of one for transmitting information.

3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm but would require clumsy restatement in terms of the number of possibilities.

The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly *bits*, a word suggested by J. W. Tukey. A device with two stable positions, such as a relay or a flip-flop circuit, can store one bit of information. N such devices can store N bits, since the total number of possible states is 2^N and $\log_2 2^N = N$. If the base 10 is used the units may be called decimal digits. Since

$$\begin{aligned}\log_2 M &= \log_{10} M / \log_{10} 2 \\ &= 3.32 \log_{10} M,\end{aligned}$$

a decimal digit is about $3\frac{1}{3}$ bits. A digit wheel on a desk computing machine has ten stable positions and therefore has a storage capacity of one decimal digit. In analytical work where integration and differentiation are involved the base e is sometimes useful. The resulting units of information will be called natural units. Change from the base a to base b merely requires multiplication by $\log_b a$.

By a communication system we will mean a system of the type indicated schematically in Fig. 1. It consists of essentially five parts:

1. An *information source* which produces a message or sequence of messages to be communicated to the receiving terminal. The message may be of various types: e.g. (a) A sequence of letters as in a telegraph or teletype system; (b) A single function of time $f(t)$ as in radio or telephony; (c) A function of time and other variables as in black and white television—here the message may be thought of as a function $f(x, y, t)$ of two space coordinates and time, the light intensity at point (x, y) and time t on a pickup tube plate; (d) Two or more functions of time, say $f(t)$, $g(t)$, $h(t)$ —this is the case in “three dimensional” sound transmission or if the system is intended to service several individual channels in multiplex; (e) Several functions of

several variables—in color television the message consists of three functions $f(x, y, t)$, $g(x, y, t)$, $h(x, y, t)$ defined in a three-dimensional continuum—we may also think of these three functions as components of a vector field defined in the region—similarly, several black and white television sources would produce “messages” consisting of a number of functions of three variables; (f) Various combinations also occur, for example in television with an associated audio channel.

2. A *transmitter* which operates on the message in some way to produce a signal suitable for transmission over the channel. In telephony this operation consists merely of changing sound pressure into a proportional electrical current. In telegraphy we have an encoding operation which produces a sequence of dots, dashes and spaces on the channel corresponding to the message. In a multiplex PCM system the different speech functions must be sampled, compressed, quantized and encoded, and finally interleaved

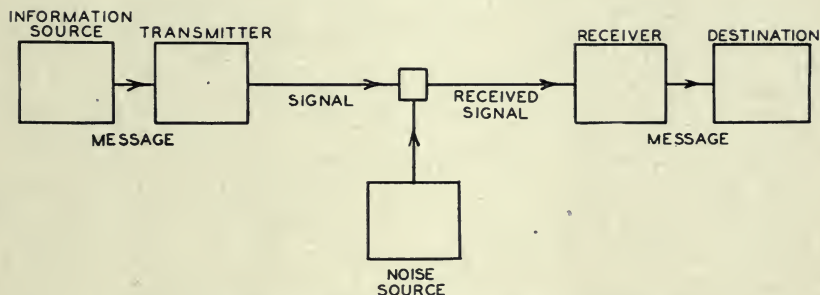


Fig. 1—Schematic diagram of a general communication system.

properly to construct the signal. Vocoder systems, television, and frequency modulation are other examples of complex operations applied to the message to obtain the signal.

3. The *channel* is merely the medium used to transmit the signal from transmitter to receiver. It may be a pair of wires, a coaxial cable, a band of radio frequencies, a beam of light, etc.

4. The *receiver* ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.

5. The *destination* is the person (or thing) for whom the message is intended.

We wish to consider certain general problems involving communication systems. To do this it is first necessary to represent the various elements involved as mathematical entities, suitably idealized from their physical counterparts. We may roughly classify communication systems into three main categories: discrete, continuous and mixed. By a discrete system we will mean one in which both the message and the signal are a sequence of

discrete symbols. A typical case is telegraphy where the message is a sequence of letters and the signal a sequence of dots, dashes and spaces. A continuous system is one in which the message and signal are both treated as continuous functions, e.g. radio or television. A mixed system is one in which both discrete and continuous variables appear, e.g., PCM transmission of speech.

We first consider the discrete case. This case has applications not only in communication theory, but also in the theory of computing machines, the design of telephone exchanges and other fields. In addition the discrete case forms a foundation for the continuous and mixed cases which will be treated in the second half of the paper.

PART I: DISCRETE NOISELESS SYSTEMS

1. THE DISCRETE NOISELESS CHANNEL

Teletype and telegraphy are two simple examples of a discrete channel for transmitting information. Generally, a discrete channel will mean a system whereby a sequence of choices from a finite set of elementary symbols $S_1 \cdots S_n$ can be transmitted from one point to another. Each of the symbols S_i is assumed to have a certain duration in time t_i seconds (not necessarily the same for different S_i , for example the dots and dashes in telegraphy). It is not required that all possible sequences of the S_i be capable of transmission on the system; certain sequences only may be allowed. These will be possible signals for the channel. Thus in telegraphy suppose the symbols are: (1) A dot, consisting of line closure for a unit of time and then line open for a unit of time; (2) A dash, consisting of three time units of closure and one unit open; (3) A letter space consisting of, say, three units of line open; (4) A word space of six units of line open. We might place the restriction on allowable sequences that no spaces follow each other (for if two letter spaces are adjacent, it is identical with a word space). The question we now consider is how one can measure the capacity of such a channel to transmit information.

In the teletype case where all symbols are of the same duration, and any sequence of the 32 symbols is allowed the answer is easy. Each symbol represents five bits of information. If the system transmits n symbols per second it is natural to say that the channel has a capacity of $5n$ bits per second. This does not mean that the teletype channel will always be transmitting information at this rate—this is the maximum possible rate and whether or not the actual rate reaches this maximum depends on the source of information which feeds the channel, as will appear later.

In the more general case with different lengths of symbols and constraints on the allowed sequences, we make the following definition:

Definition: The capacity C of a discrete channel is given by

$$C = \lim_{T \rightarrow \infty} \frac{\log N(T)}{T}$$

where $N(T)$ is the number of allowed signals of duration T .

It is easily seen that in the teletype case this reduces to the previous result. It can be shown that the limit in question will exist as a finite number in most cases of interest. Suppose all sequences of the symbols $S_1, \dots, S_n$ are allowed and these symbols have durations $t_1, \dots, t_n$. What is the channel capacity? If $N(t)$ represents the number of sequences of duration t we have

$$N(t) = N(t - t_1) + N(t - t_2) + \dots + N(t - t_n)$$

The total number is equal to the sum of the numbers of sequences ending in $S_1, S_2, \dots, S_n$ and these are $N(t - t_1), N(t - t_2), \dots, N(t - t_n)$, respectively. According to a well known result in finite differences, $N(t)$ is then asymptotic for large t to X_0^t where X_0 is the largest real solution of the characteristic equation:

$$X^{-t_1} + X^{-t_2} + \dots + X^{-t_n} = 1$$

and therefore

$$C = \log X_0$$

In case there are restrictions on allowed sequences we may still often obtain a difference equation of this type and find C from the characteristic equation. In the telegraph case mentioned above

$$N(t) = N(t - 2) + N(t - 4) + N(t - 5) + N(t - 7) + N(t - 8) \\ + N(t - 10)$$

as we see by counting sequences of symbols according to the last or next to the last symbol occurring. Hence C is $-\log \mu_0$ where μ_0 is the positive root of $1 = \mu^2 + \mu^4 + \mu^5 + \mu^7 + \mu^8 + \mu^{10}$. Solving this we find $C = 0.539$.

A very general type of restriction which may be placed on allowed sequences is the following: We imagine a number of possible states $a_1, a_2, \dots, a_m$. For each state only certain symbols from the set $S_1, \dots, S_n$ can be transmitted (different subsets for the different states). When one of these has been transmitted the state changes to a new state depending both on the old state and the particular symbol transmitted. The telegraph case is a simple example of this. There are two states depending on whether or not

a space was the last symbol transmitted. If so then only a dot or a dash can be sent next and the state always changes. If not, any symbol can be transmitted and the state changes if a space is sent, otherwise it remains the same. The conditions can be indicated in a linear graph as shown in Fig. 2. The junction points correspond to the states and the lines indicate the symbols possible in a state and the resulting state. In Appendix I it is shown that if the conditions on allowed sequences can be described in this form C will exist and can be calculated in accordance with the following result:

Theorem 1: Let $b_{ij}^{(s)}$ be the duration of the s^{th} symbol which is allowable in state i and leads to state j . Then the channel capacity C is equal to $\log W$ where W is the largest real root of the determinant equation:

$$\left| \sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right| = 0,$$

where $\delta_{ij} = 1$ if $i = j$ and is zero otherwise.

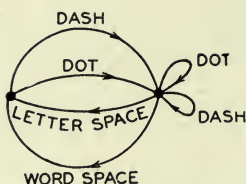


Fig. 2—Graphical representation of the constraints on telegraph symbols.

For example, in the telegraph case (Fig. 2) the determinant is:

$$\begin{vmatrix} -1 & (W^{-2} + W^{-4}) \\ (W^{-3} + W^{-6}) & (W^{-2} + W^{-4} - 1) \end{vmatrix} = 0$$

On expansion this leads to the equation given above for this case.

2. THE DISCRETE SOURCE OF INFORMATION

We have seen that under very general conditions the logarithm of the number of possible signals in a discrete channel increases linearly with time. The capacity to transmit information can be specified by giving this rate of increase, the number of bits per second required to specify the particular signal used.

We now consider the information source. How is an information source to be described mathematically, and how much information in bits per second is produced in a given source? The main point at issue is the effect of statistical knowledge about the source in reducing the required capacity

of the channel, by the use of proper encoding of the information. In telegraphy, for example, the messages to be transmitted consist of sequences of letters. These sequences, however, are not completely random. In general, they form sentences and have the statistical structure of, say, English. The letter E occurs more frequently than Q, the sequence TH more frequently than XP, etc. The existence of this structure allows one to make a saving in time (or channel capacity) by properly encoding the message sequences into signal sequences. This is already done to a limited extent in telegraphy by using the shortest channel symbol, a dot, for the most common English letter E; while the infrequent letters, Q, X, Z are represented by longer sequences of dots and dashes. This idea is carried still further in certain commercial codes where common words and phrases are represented by four- or five-letter code groups with a considerable saving in average time. The standardized greeting and anniversary telegrams now in use extend this to the point of encoding a sentence or two into a relatively short sequence of numbers.

We can think of a discrete source as generating the message, symbol by symbol. It will choose successive symbols according to certain probabilities depending, in general, on preceding choices as well as the particular symbols in question. A physical system, or a mathematical model of a system which produces such a sequence of symbols governed by a set of probabilities is known as a stochastic process.³ We may consider a discrete source, therefore, to be represented by a stochastic process. Conversely, any stochastic process which produces a discrete sequence of symbols chosen from a finite set may be considered a discrete source. This will include such cases as:

1. Natural written languages such as English, German, Chinese.
2. Continuous information sources that have been rendered discrete by some quantizing process. For example, the quantized speech from a PCM transmitter, or a quantized television signal.
3. Mathematical cases where we merely define abstractly a stochastic process which generates a sequence of symbols. The following are examples of this last type of source.

(A) Suppose we have five letters A, B, C, D, E which are chosen each with probability .2, successive choices being independent. This would lead to a sequence of which the following is a typical example.

B D C B C E C C C A D C B D D A A E C E E A
A B B D A E E C A C E E B A E E C B C E A D

This was constructed with the use of a table of random numbers.⁴

³ See, for example, S. Chandrasekhar, "Stochastic Problems in Physics and Astronomy," *Reviews of Modern Physics*, v. 15, No. 1, January 1943, p. 1.

⁴ Kendall and Smith, "Tables of Random Sampling Numbers," Cambridge, 1939.

- (B) Using the same five letters let the probabilities be .4, .1, .2, .2, .1 respectively, with successive choices independent. A typical message from this source is then:

A A A C D C B D C E A A D A D A C E D A

E A D C A B E D A D D C E C A A A A A D

- (C) A more complicated structure is obtained if successive symbols are not chosen independently but their probabilities depend on preceding letters. In the simplest case of this type a choice depends only on the preceding letter and not on ones before that. The statistical structure can then be described by a set of transition probabilities $p_i(j)$, the probability that letter i is followed by letter j . The indices i and j range over all the possible symbols. A second equivalent way of specifying the structure is to give the "digram" probabilities $p(i, j)$, i.e., the relative frequency of the digram ij . The letter frequencies $p(i)$, (the probability of letter i), the transition probabilities $p_i(j)$ and the digram probabilities $p(i, j)$ are related by the following formulas.

$$p(i) = \sum_j p(i, j) = \sum_j p(j, i) = \sum_j p(j) p_j(i)$$

$$p(i, j) = p(i) p_i(j)$$

$$\sum_j p_i(j) = \sum_i p(i) = \sum_{i,j} p(i, j) = 1.$$

As a specific example suppose there are three letters A, B, C with the probability tables:

$p_i(j)$	j			i	$p(i)$	$p(i, j)$	j		
	A	B	C				A	B	C
A	0	$\frac{4}{5}$	$\frac{1}{5}$	A	$\frac{9}{27}$	A	0	$\frac{4}{15}$	$\frac{1}{15}$
B	$\frac{1}{2}$	$\frac{1}{2}$	0	B	$\frac{16}{27}$	B	$\frac{8}{27}$	$\frac{8}{27}$	0
C	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{1}{10}$	C	$\frac{2}{27}$	C	$\frac{1}{27}$	$\frac{4}{135}$	$\frac{1}{135}$

A typical message from this source is the following:

A B B A B A B A B A B A B A B B B A B B B B A B

A B A B A B A B B B A C A C A B B A B B B A B B

A B A C B B B A B A

The next increase in complexity would involve trigram frequencies but no more. The choice of a letter would depend on the preceding two letters but not on the message before that point. A set of trigram frequencies $p(i, j, k)$ or equivalently a set of transition prob-

abilities $p_{ij}(k)$ would be required. Continuing in this way one obtains successively more complicated stochastic processes. In the general n -gram case a set of n -gram probabilities $p(i_1, i_2, \dots, i_n)$ or of transition probabilities $p_{i_1, i_2, \dots, i_{n-1}}(i_n)$ is required to specify the statistical structure.

- (D) Stochastic processes can also be defined which produce a text consisting of a sequence of "words." Suppose there are five letters A, B, C, D, E and 16 "words" in the language with associated probabilities:

.10 A	.16 BEBE	.11 CABED	.04 DEB
.04 ADEB	.04 BED	.05 CEED	.15 DEED
.05 ADEE	.02 BEED	.08 DAB	.01 EAB
.01 BADD	.05 CA	.04 DAD	.05 EE

Suppose successive "words" are chosen independently and are separated by a space. A typical message might be:

DAB EE A BEBE DEED DEB ADEE ADEE EE DEB BEBE
BEBE BEBE ADEE BED DEED DEED CEED ADEE A DEED
DEED BEBE CABED BEBE BED DAB DEED ADEB

If all the words are of finite length this process is equivalent to one of the preceding type, but the description may be simpler in terms of the word structure and probabilities. We may also generalize here and introduce transition probabilities between words, etc.

These artificial languages are useful in constructing simple problems and examples to illustrate various possibilities. We can also approximate to a natural language by means of a series of simple artificial languages. The zero-order approximation is obtained by choosing all letters with the same probability and independently. The first-order approximation is obtained by choosing successive letters independently but each letter having the same probability that it does in the natural language.⁵ Thus, in the first-order approximation to English, E is chosen with probability .12 (its frequency in normal English) and W with probability .02, but there is no influence between adjacent letters and no tendency to form the preferred digrams such as *TH*, *ED*, etc. In the second-order approximation, digram structure is introduced. After a letter is chosen, the next one is chosen in accordance with the frequencies with which the various letters follow the first one. This requires a table of digram frequencies $p_i(j)$. In the third-order approximation, trigram structure is introduced. Each letter is chosen with probabilities which depend on the preceding two letters.

⁵ Letter, digram and trigram frequencies are given in "Secret and Urgent" by Fletcher Pratt, Blue Ribbon Books 1939. Word frequencies are tabulated in "Relative Frequency of English Speech Sounds," G. Dewey, Harvard University Press, 1923.

3. THE SERIES OF APPROXIMATIONS TO ENGLISH

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol "alphabet," the 26 letters and a space.

1. Zero-order approximation (symbols independent and equi-probable).
XFOML RXKHRJFFJUJ ZLPWCFWKCYJ
FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD

2. First-order approximation (symbols independent but with frequencies of English text).

OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI
ALHENHTTPA OOBTTVA NAH BRL

3. Second-order approximation (digram structure as in English).
ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY
ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO
TIZIN ANDY TOBE SEACE CTISBE

4. Third-order approximation (trigram structure as in English).
IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID
PONDENOME OF DEMONSTURES OF THE REPTAGIN IS
REGOACTIONA OF CRE

5. First-Order Word Approximation. Rather than continue with tetragram, $\dots$, n -gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

REPRESENTING AND SPEEDILY IS AN GOOD APT OR
COME CAN DIFFERENT NATURAL HERE HE THE A IN
CAME THE TO OF TO EXPERT GRAY COME TO FUR-
NISHES THE LINE MESSAGE HAD BE THESE.

6. Second-Order Word Approximation. The word transition probabilities are correct but no further structure is included.

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH
WRITER THAT THE CHARACTER OF THIS POINT IS
THEREFORE ANOTHER METHOD FOR THE LETTERS
THAT THE TIME OF WHO EVER TOLD THE PROBLEM
FOR AN UNEXPECTED

The resemblance to ordinary English text increases quite noticeably at each of the above steps. Note that these samples have reasonably good structure out to about twice the range that is taken into account in their construction. Thus in (3) the statistical process insures reasonable text for two-letter sequence, but four-letter sequences from the sample can usually be fitted into good sentences. In (6) sequences of four or more

words can easily be placed in sentences without unusual or strained constructions. The particular sequence of ten words "attack on an English writer that the character of this" is not at all unreasonable. It appears then that a sufficiently complex stochastic process will give a satisfactory representation of a discrete source.

The first two samples were constructed by the use of a book of random numbers in conjunction with (for example 2) a table of letter frequencies. This method might have been continued for (3), (4), and (5), since digram, trigram, and word frequency tables are available, but a simpler equivalent method was used. To construct (3) for example, one opens a book at random and selects a letter at random on the page. This letter is recorded. The book is then opened to another page and one reads until this letter is encountered. The succeeding letter is then recorded. Turning to another page this second letter is searched for and the succeeding letter recorded, etc. A similar process was used for (4), (5), and (6). It would be interesting if further approximations could be constructed, but the labor involved becomes enormous at the next stage.

4. GRAPHICAL REPRESENTATION OF A MARKOFF PROCESS

Stochastic processes of the type described above are known mathematically as discrete Markoff processes and have been extensively studied in the literature.⁶ The general case can be described as follows: There exist a finite number of possible "states" of a system; $S_1, S_2, \dots, S_n$. In addition there is a set of transition probabilities; $p_i(j)$ the probability that if the system is in state S_i it will next go to state S_j . To make this Markoff process into an information source we need only assume that a letter is produced for each transition from one state to another. The states will correspond to the "residue of influence" from preceding letters.

The situation can be represented graphically as shown in Figs. 3, 4 and 5. The "states" are the junction points in the graph and the probabilities and letters produced for a transition are given beside the corresponding line. Figure 3 is for the example B in Section 2, while Fig. 4 corresponds to the example C. In Fig. 3 there is only one state since successive letters are independent. In Fig. 4 there are as many states as letters. If a trigram example were constructed there would be at most n^2 states corresponding to the possible pairs of letters preceding the one being chosen. Figure 5 is a graph for the case of word structure in example D. Here S corresponds to the "space" symbol.

⁶ For a detailed treatment see M. Frechet, "Methods des fonctions arbitraires. Theorie des enenements en chaine dans le cas d'un nombre fini d'etats possibles." Paris, Gauthier-Villars, 1938.

5. ERGODIC AND MIXED SOURCES

As we have indicated above a discrete source for our purposes can be considered to be represented by a Markoff process. Among the possible discrete Markoff processes there is a group with special properties of significance in

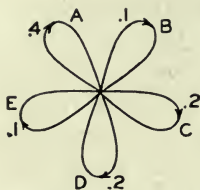


Fig. 3—A graph corresponding to the source in example B.

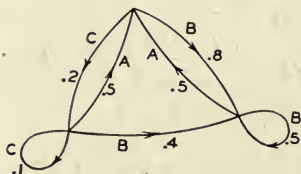


Fig. 4—A graph corresponding to the source in example C.

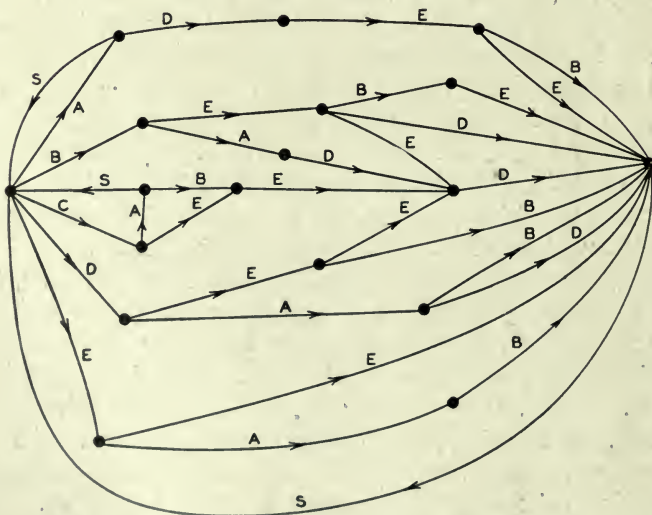


Fig. 5—A graph corresponding to the source in example D.

communication theory. This special class consists of the "ergodic" processes and we shall call the corresponding sources ergodic sources. Although a rigorous definition of an ergodic process is somewhat involved, the general idea is simple. In an ergodic process every sequence produced by the proc-

ess is the same in statistical properties. Thus the letter frequencies, digram frequencies, etc., obtained from particular sequences will, as the lengths of the sequences increase, approach definite limits independent of the particular sequence. Actually this is not true of every sequence but the set for which it is false has probability zero. Roughly the ergodic property means statistical homogeneity.

All the examples of artificial languages given above are ergodic. This property is related to the structure of the corresponding graph. If the graph has the following two properties⁷ the corresponding process will be ergodic:

1. The graph does not consist of two isolated parts A and B such that it is impossible to go from junction points in part A to junction points in part B along lines of the graph in the direction of arrows and also impossible to go from junctions in part B to junctions in part A.
2. A closed series of lines in the graph with all arrows on the lines pointing in the same orientation will be called a "circuit." The "length" of a circuit is the number of lines in it. Thus in Fig. 5 the series BEBES is a circuit of length 5. The second property required is that the greatest common divisor of the lengths of all circuits in the graph be one.

If the first condition is satisfied but the second one violated by having the greatest common divisor equal to $d > 1$, the sequences have a certain type of periodic structure. The various sequences fall into d different classes which are statistically the same apart from a shift of the origin (i.e., which letter in the sequence is called letter 1). By a shift of from 0 up to $d - 1$ any sequence can be made statistically equivalent to any other. A simple example with $d = 2$ is the following: There are three possible letters a, b, c . Letter a is followed with either b or c with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ respectively. Either b or c is always followed by letter a . Thus a typical sequence is

a b a c a c a c a b a c a b a b a c a c

This type of situation is not of much importance for our work.

If the first condition is violated the graph may be separated into a set of subgraphs each of which satisfies the first condition. We will assume that the second condition is also satisfied for each subgraph. We have in this case what may be called a "mixed" source made up of a number of pure components. The components correspond to the various subgraphs. If $L_1, L_2, L_3, \dots$ are the component sources we may write

$$L = p_1 L_1 + p_2 L_2 + p_3 L_3 + \dots$$

where p_i is the probability of the component source L_i .

⁷ These are restatements in terms of the graph of conditions given in Frechet.

Physically the situation represented is this: There are several different sources $L_1, L_2, L_3, \dots$ which are each of homogeneous statistical structure (i.e., they are ergodic). We do not know *a priori* which is to be used, but once the sequence starts in a given pure component L_i it continues indefinitely according to the statistical structure of that component.

As an example one may take two of the processes defined above and assume $p_1 = .2$ and $p_2 = .8$. A sequence from the mixed source

$$L = .2 L_1 + .8 L_2$$

would be obtained by choosing first L_1 or L_2 with probabilities .2 and .8 and after this choice generating a sequence from whichever was chosen.

Except when the contrary is stated we shall assume a source to be ergodic. This assumption enables one to identify averages along a sequence with averages over the ensemble of possible sequences (the probability of a discrepancy being zero). For example the relative frequency of the letter A in a particular infinite sequence will be, with probability one, equal to its relative frequency in the ensemble of sequences.

If P_i is the probability of state i and $p_i(j)$ the transition probability to state j , then for the process to be stationary it is clear that the P_i must satisfy equilibrium conditions:

$$P_j = \sum_i P_i p_i(j).$$

In the ergodic case it can be shown that with any starting conditions the probabilities $P_j(N)$ of being in state j after N symbols, approach the equilibrium values as $N \rightarrow \infty$.

6. CHOICE, UNCERTAINTY AND ENTROPY

We have represented a discrete information source as a Markoff process. Can we define a quantity which will measure, in some sense, how much information is "produced" by such a process, or better, at what rate information is produced?

Suppose we have a set of possible events whose probabilities of occurrence are $p_1, p_2, \dots, p_n$. These probabilities are known but that is all we know concerning which event will occur. Can we find a measure of how much "choice" is involved in the selection of the event or of how uncertain we are of the outcome?

If there is such a measure, say $H(p_1, p_2, \dots, p_n)$, it is reasonable to require of it the following properties:

1. H should be continuous in the p_i .
2. If all the p_i are equal, $p_i = \frac{1}{n}$, then H should be a monotonic increasing

function of n . With equally likely events there is more choice, or uncertainty, when there are more possible events.

3. If a choice be broken down into two successive choices, the original H should be the weighted sum of the individual values of H . The meaning of this is illustrated in Fig. 6. At the left we have three possibilities $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, $p_3 = \frac{1}{6}$. On the right we first choose between two possibilities each with probability $\frac{1}{2}$, and if the second occurs make another choice with probabilities $\frac{2}{3}$, $\frac{1}{3}$. The final results have the same probabilities as before. We require, in this special case, that

$$H(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}) = H(\frac{1}{2}, \frac{1}{2}) + \frac{1}{2}H(\frac{2}{3}, \frac{1}{3})$$

The coefficient $\frac{1}{2}$ is because this second choice only occurs half the time.

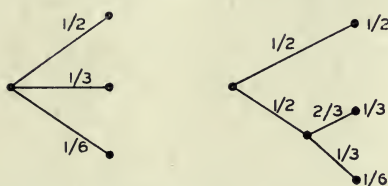


Fig. 6—Decomposition of a choice from three possibilities.

In Appendix II, the following result is established:

Theorem 2: The only H satisfying the three above assumptions is of the form:

$$H = -K \sum_{i=1}^n p_i \log p_i$$

where K is a positive constant.

This theorem, and the assumptions required for its proof, are in no way necessary for the present theory. It is given chiefly to lend a certain plausibility to some of our later definitions. The real justification of these definitions, however, will reside in their implications.

Quantities of the form $H = -\sum p_i \log p_i$ (the constant K merely amounts to a choice of a unit of measure) play a central role in information theory as measures of information, choice and uncertainty. The form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics⁸ where p_i is the probability of a system being in cell i of its phase space. H is then, for example, the H in Boltzmann's famous H theorem. We shall call $H = -\sum p_i \log p_i$ the entropy of the set of probabilities

⁸ See, for example, R. C. Tolman, "Principles of Statistical Mechanics," Oxford, Clarendon, 1938.

$p_1, \dots, p_n$. If x is a chance variable we will write $H(x)$ for its entropy; thus x is not an argument of a function but a label for a number, to differentiate it from $H(y)$ say, the entropy of the chance variable y .

The entropy in the case of two possibilities with probabilities p and $q = 1 - p$, namely

$$H = -(p \log p + q \log q)$$

is plotted in Fig. 7 as a function of p .

The quantity H has a number of interesting properties which further substantiate it as a reasonable measure of choice or information.

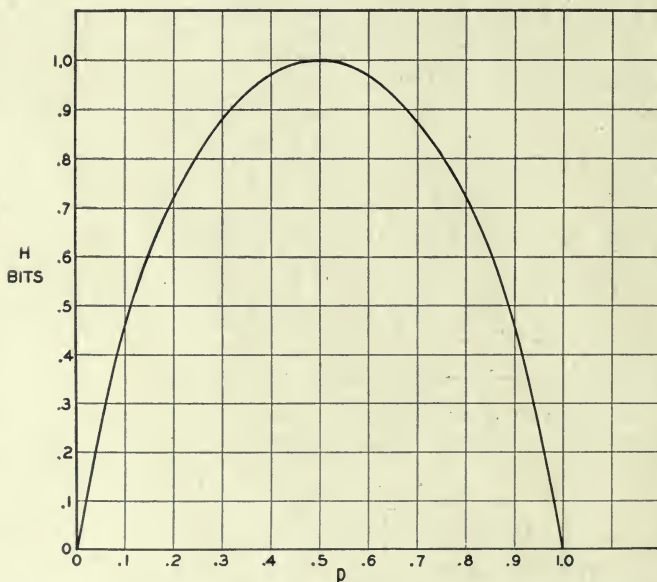


Fig. 7—Entropy in the case of two possibilities with probabilities p and $(1 - p)$.

1. $H = 0$ if and only if all the p_i but one are zero, this one having the value unity. Thus only when we are certain of the outcome does H vanish. Otherwise H is positive.

2. For a given n , H is a maximum and equal to $\log n$ when all the p_i are equal (i.e., $\frac{1}{n}$). This is also intuitively the most uncertain situation.

3. Suppose there are two events, x and y , in question with m possibilities for the first and n for the second. Let $p(i, j)$ be the probability of the joint occurrence of i for the first and j for the second. The entropy of the joint event is

$$H(x, y) = - \sum_{i,j} p(i, j) \log p(i, j)$$

while

$$H(x) = - \sum_{i,j} p(i, j) \log \sum_j p(i, j)$$

$$H(y) = - \sum_{i,j} p(i, j) \log \sum_i p(i, j).$$

It is easily shown that

$$H(x, y) \leq H(x) + H(y)$$

with equality only if the events are independent (i.e., $p(i, j) = p(i) p(j)$). The uncertainty of a joint event is less than or equal to the sum of the individual uncertainties.

4. Any change toward equalization of the probabilities $p_1, p_2, \dots, p_n$ increases H . Thus if $p_1 < p_2$ and we increase p_1 , decreasing p_2 an equal amount so that p_1 and p_2 are more nearly equal, then H increases. More generally, if we perform any "averaging" operation on the p_i of the form

$$p'_i = \sum_j a_{ij} p_j$$

where $\sum_i a_{ij} = \sum_j a_{ij} = 1$, and all $a_{ij} \geq 0$, then H increases (except in the special case where this transformation amounts to no more than a permutation of the p_j with H of course remaining the same).

5. Suppose there are two chance events x and y as in 3, not necessarily independent. For any particular value i that x can assume there is a conditional probability $p_i(j)$ that y has the value j . This is given by

$$p_i(j) = \frac{p(i, j)}{\sum_j p(i, j)}.$$

We define the *conditional entropy* of y , $H_x(y)$ as the average of the entropy of y for each value of x , weighted according to the probability of getting that particular x . That is

$$H_x(y) = - \sum_{i,j} p(i, j) \log p_i(j).$$

This quantity measures how uncertain we are of y on the average when we know x . Substituting the value of $p_i(j)$ we obtain

$$H_x(y) = - \sum_{i,j} p(i, j) \log p(i, j) + \sum_{i,j} p(i, j) \log \sum_j p(i, j)$$

$$= H(x, y) - H(x)$$

or

$$H(x, y) = H(x) + H_x(y)$$

The uncertainty (or entropy) of the joint event x, y is the uncertainty of x plus the uncertainty of y when x is known.

6. From 3 and 5 we have

$$H(x) + H(y) \geq H(x, y) = H(x) + H_x(y)$$

Hence

$$H(y) \geq H_x(y)$$

The uncertainty of y is never increased by knowledge of x . It will be decreased unless x and y are independent events, in which case it is not changed.

7. THE ENTROPY OF AN INFORMATION SOURCE

Consider a discrete source of the finite state type considered above. For each possible state i there will be a set of probabilities $p_i(j)$ of producing the various possible symbols j . Thus there is an entropy H_i for each state. The entropy of the source will be defined as the average of these H_i weighted in accordance with the probability of occurrence of the states in question:

$$\begin{aligned} H &= \sum_i P_i H_i \\ &= - \sum_{i,j} P_i p_i(j) \log p_i(j) \end{aligned}$$

This is the entropy of the source per symbol of text. If the Markoff process is proceeding at a definite time rate there is also an entropy per second

$$H' = \sum_i f_i H_i$$

where f_i is the average frequency (occurrences per second) of state i . Clearly

$$H' = mH$$

where m is the average number of symbols produced per second. H or H' measures the amount of information generated by the source per symbol or per second. If the logarithmic base is 2, they will represent bits per symbol or per second.

If successive symbols are independent then H is simply $-\sum p_i \log p_i$ where p_i is the probability of symbol i . Suppose in this case we consider a long message of N symbols. It will contain with high probability about $p_1 N$ occurrences of the first symbol, $p_2 N$ occurrences of the second, etc. Hence the probability of this particular message will be roughly

$$p = p_1^{p_1 N} p_2^{p_2 N} \cdots p_n^{p_n N}$$

or

$$\log p \doteq N \sum_i p_i \log p_i$$

$$\log p \doteq -NH$$

$$H \doteq \frac{\log 1/p}{N}$$

H is thus approximately the logarithm of the reciprocal probability of a typical long sequence divided by the number of symbols in the sequence. The same result holds for any source. Stated more precisely we have (see Appendix III):

Theorem 3: Given any $\epsilon > 0$ and $\delta > 0$, we can find an N_0 such that the sequences of any length $N \geq N_0$ fall into two classes:

1. A set whose total probability is less than ϵ .
2. The remainder, all of whose members have probabilities satisfying the inequality

$$\left| \frac{\log p^{-1}}{N} - H \right| < \delta$$

In other words we are almost certain to have $\frac{\log p^{-1}}{N}$ very close to H when N is large.

A closely related result deals with the number of sequences of various probabilities. Consider again the sequences of length N and let them be arranged in order of decreasing probability. We define $n(q)$ to be the number we must take from this set starting with the most probable one in order to accumulate a total probability q for those taken.

Theorem 4:

$$\lim_{N \rightarrow \infty} \frac{\log n(q)}{N} = H$$

when q does not equal 0 or 1.

We may interpret $\log n(q)$ as the number of bits required to specify the sequence when we consider only the most probable sequences with a total probability q . Then $\frac{\log n(q)}{N}$ is the number of bits per symbol for the specification. The theorem says that for large N this will be independent of q and equal to H . The rate of growth of the logarithm of the number of reasonably probable sequences is given by H , regardless of our interpretation of "reasonably probable." Due to these results, which are proved in appendix III, it is possible for most purposes to treat the long sequences as though there were just 2^{HN} of them, each with a probability 2^{-HN} .

The next two theorems show that H and H' can be determined by limiting operations directly from the statistics of the message sequences, without reference to the states and transition probabilities between states.

Theorem 5: Let $p(B_i)$ be the probability of a sequence B_i of symbols from the source. Let

$$G_N = -\frac{1}{N} \sum_i p(B_i) \log p(B_i)$$

where the sum is over all sequences B_i containing N symbols. Then G_N is a monotonic decreasing function of N and

$$\lim_{N \rightarrow \infty} G_N = H.$$

Theorem 6: Let $p(B_i, S_j)$ be the probability of sequence B_i followed by symbol S_j and $p_{B_i}(S_j) = p(B_i, S_j)/p(B_i)$ be the conditional probability of S_j after B_i . Let

$$F_N = -\sum_{i,j} p(B_i, S_j) \log p_{B_i}(S_j)$$

where the sum is over all blocks B_i of $N - 1$ symbols and over all symbols S_j . Then F_N is a monotonic decreasing function of N ,

$$F_N = NG_N - (N - 1) G_{N-1},$$

$$G_N = \frac{1}{N} \sum_1^N F_N,$$

$$F_N \leq G_N,$$

and $\lim_{N \rightarrow \infty} F_N = H$.

These results are derived in appendix III. They show that a series of approximations to H can be obtained by considering only the statistical structure of the sequences extending over 1, 2, $\dots$ N symbols. F_N is the better approximation. In fact F_N is the entropy of the N^{th} order approximation to the source of the type discussed above. If there are no statistical influences extending over more than N symbols, that is if the conditional probability of the next symbol knowing the preceding $(N - 1)$ is not changed by a knowledge of any before that, then $F_N = H$. F_N of course is the conditional entropy of the next symbol when the $(N - 1)$ preceding ones are known, while G_N is the entropy per symbol of blocks of N symbols.

The ratio of the entropy of a source to the maximum value it could have while still restricted to the same symbols will be called its *relative entropy*. This is the maximum compression possible when we encode into the same alphabet. One minus the relative entropy is the *redundancy*. The redun-

dancy of ordinary English, not considering statistical structure over greater distances than about eight letters is roughly 50%. This means that when we write English half of what we write is determined by the structure of the language and half is chosen freely. The figure 50% was found by several independent methods which all gave results in this neighborhood. One is by calculation of the entropy of the approximations to English. A second method is to delete a certain fraction of the letters from a sample of English text and then let someone attempt to restore them. If they can be restored when 50% are deleted the redundancy must be greater than 50%. A third method depends on certain known results in cryptography.

Two extremes of redundancy in English prose are represented by Basic English and by James Joyces' book "Finigans Wake." The Basic English vocabulary is limited to 850 words and the redundancy is very high. This is reflected in the expansion that occurs when a passage is translated into Basic English. Joyce on the other hand enlarges the vocabulary and is alleged to achieve a compression of semantic content.

The redundancy of a language is related to the existence of crossword puzzles. If the redundancy is zero any sequence of letters is a reasonable text in the language and any two dimensional array of letters forms a crossword puzzle. If the redundancy is too high the language imposes too many constraints for large crossword puzzles to be possible. A more detailed analysis shows that if we assume the constraints imposed by the language are of a rather chaotic and random nature, large crossword puzzles are just possible when the redundancy is 50%. If the redundancy is 33%, three dimensional crossword puzzles should be possible, etc.

8. REPRESENTATION OF THE ENCODING AND DECODING OPERATIONS

We have yet to represent mathematically the operations performed by the transmitter and receiver in encoding and decoding the information. Either of these will be called a discrete transducer. The input to the transducer is a sequence of input symbols and its output a sequence of output symbols. The transducer may have an internal memory so that its output depends not only on the present input symbol but also on the past history. We assume that the internal memory is finite, i.e. there exists a finite number m of possible states of the transducer and that its output is a function of the present state and the present input symbol. The next state will be a second function of these two quantities. Thus a transducer can be described by two functions:

$$\begin{aligned} y_n &= f(x_n, \alpha_n) \\ \alpha_{n+1} &= g(x_n, \alpha_n) \end{aligned}$$

where: x_n is the n^{th} input symbol,

α_n is the state of the transducer when the n^{th} input symbol is introduced,

y_n is the output symbol (or sequence of output symbols) produced when

x_n is introduced if the state is α_n .

If the output symbols of one transducer can be identified with the input symbols of a second, they can be connected in tandem and the result is also a transducer. If there exists a second transducer which operates on the output of the first and recovers the original input, the first transducer will be called non-singular and the second will be called its inverse.

Theorem 7: The output of a finite state transducer driven by a finite state statistical source is a finite state statistical source, with entropy (per unit time) less than or equal to that of the input. If the transducer is non-singular they are equal.

Let α represent the state of the source, which produces a sequence of symbols x_i ; and let β be the state of the transducer, which produces, in its output, blocks of symbols y_j . The combined system can be represented by the "product state space" of pairs (α, β) . Two points in the space, (α_1, β_1) and (α_2, β_2) , are connected by a line if α_1 can produce an x which changes β_1 to β_2 , and this line is given the probability of that x in this case. The line is labeled with the block of y_j symbols produced by the transducer. The entropy of the output can be calculated as the weighted sum over the states. If we sum first on β each resulting term is less than or equal to the corresponding term for α , hence the entropy is not increased. If the transducer is non-singular let its output be connected to the inverse transducer. If H'_1 , H'_2 and H'_3 are the output entropies of the source, the first and second transducers respectively, then $H'_1 \geq H'_2 \geq H'_3 = H'_1$ and therefore $H'_1 = H'_2$.

Suppose we have a system of constraints on possible sequences of the type which can be represented by a linear graph as in Fig. 2. If probabilities $p_{ij}^{(s)}$ were assigned to the various lines connecting state i to state j this would become a source. There is one particular assignment which maximizes the resulting entropy (see Appendix IV).

Theorem 8: Let the system of constraints considered as a channel have a capacity C . If we assign

$$p_{ij}^{(s)} = \frac{B_j}{B_i} C^{-\ell_{ij}^{(s)}}$$

where $\ell_{ij}^{(s)}$ is the duration of the s^{th} symbol leading from state i to state j and the B_i satisfy

$$B_i = \sum_{s,j} B_j C^{-\ell_{ij}^{(s)}}$$

then H is maximized and equal to C .

By proper assignment of the transition probabilities the entropy of symbols on a channel can be maximized at the channel capacity.

9. THE FUNDAMENTAL THEOREM FOR A NOISELESS CHANNEL

We will now justify our interpretation of H as the rate of generating information by proving that H determines the channel capacity required with most efficient coding.

Theorem 9: Let a source have entropy H (bits per symbol) and a channel have a capacity C (bits per second). Then it is possible to encode the output of the source in such a way as to transmit at the average rate $\frac{C}{H} - \epsilon$ symbols per second over the channel where ϵ is arbitrarily small. It is not possible to transmit at an average rate greater than $\frac{C}{H}$.

The converse part of the theorem, that $\frac{C}{H}$ cannot be exceeded, may be proved by noting that the entropy of the channel input per second is equal to that of the source, since the transmitter must be non-singular, and also this entropy cannot exceed the channel capacity. Hence $H' \leq C$ and the number of symbols per second $= H'/H \leq C/H$.

The first part of the theorem will be proved in two different ways. The first method is to consider the set of all sequences of N symbols produced by the source. For N large we can divide these into two groups, one containing less than $2^{(H+\eta)N}$ members and the second containing less than 2^{RN} members (where R is the logarithm of the number of different symbols) and having a total probability less than μ . As N increases η and μ approach zero. The number of signals of duration T in the channel is greater than $2^{(C-\theta)T}$ with θ small when T is large. If we choose

$$T = \left(\frac{H}{C} + \lambda \right) N,$$

then there will be a sufficient number of sequences of channel symbols for the high probability group when N and T are sufficiently large (however small λ) and also some additional ones. The high probability group is coded in an arbitrary one to one way into this set. The remaining sequences are represented by larger sequences, starting and ending with one of the sequences not used for the high probability group. This special sequence acts as a start and stop signal for a different code. In between a sufficient time is allowed to give enough different sequences for all the low probability messages. This will require

$$T_1 = \left(\frac{R}{C} + \varphi \right) N$$

where φ is small. The mean rate of transmission in message symbols per second will then be greater than

$$\left[(1 - \delta) \frac{T}{N} + \delta \frac{T_1}{N} \right]^{-1} = \left[(1 - \delta) \left(\frac{H}{C} + \lambda \right) + \delta \left(\frac{R}{C} + \varphi \right) \right]^{-1}$$

As N increases δ , λ and φ approach zero and the rate approaches $\frac{C}{H}$.

Another method of performing this coding and proving the theorem can be described as follows: Arrange the messages of length N in order of decreasing probability and suppose their probabilities are $p_1 \geq p_2 \geq p_3 \dots \geq p_n$.

Let $P_s = \sum_{i=1}^{s-1} p_i$; that is P_s is the cumulative probability up to, but not including, p_s . We first encode into a binary system. The binary code for message s is obtained by expanding P_s as a binary number. The expansion is carried out to m_s places, where m_s is the integer satisfying:

$$\log_2 \frac{1}{p_s} \leq m_s < 1 + \log_2 \frac{1}{p_s}$$

Thus the messages of high probability are represented by short codes and those of low probability by long codes. From these inequalities we have

$$\frac{1}{2^{m_s}} \leq p_s < \frac{1}{2^{m_s-1}}.$$

The code for P_s will differ from all succeeding ones in one or more of its m_s places, since all the remaining P_i are at least $\frac{1}{2^{m_s}}$ larger and their binary expansions therefore differ in the first m_s places. Consequently all the codes are different and it is possible to recover the message from its code. If the channel sequences are not already sequences of binary digits, they can be ascribed binary numbers in an arbitrary fashion and the binary code thus translated into signals suitable for the channel.

The average number H' of binary digits used per symbol of original message is easily estimated. We have

$$H' = \frac{1}{N} \sum m_s p_s$$

But,

$$\frac{1}{N} \sum \left(\log_2 \frac{1}{p_s} \right) p_s \leq \frac{1}{N} \sum m_s p_s < \frac{1}{N} \sum \left(1 + \log_2 \frac{1}{p_s} \right) p_s$$

and therefore,

$$-\sum p_s \log p_s \leq H' < \frac{1}{N} - \sum p_s \log p_s$$

As N increases $-\sum p_s \log p_s$ approaches H , the entropy of the source and H' approaches H .

We see from this that the inefficiency in coding, when only a finite delay of N symbols is used, need not be greater than $\frac{1}{N}$ plus the difference between the true entropy H and the entropy G_N calculated for sequences of length N . The per cent excess time needed over the ideal is therefore less than

$$\frac{G_N}{H} + \frac{1}{HN} - 1.$$

This method of encoding is substantially the same as one found independently by R. M. Fano.⁹ His method is to arrange the messages of length N in order of decreasing probability. Divide this series into two groups of as nearly equal probability as possible. If the message is in the first group its first binary digit will be 0, otherwise 1. The groups are similarly divided into subsets of nearly equal probability and the particular subset determines the second binary digit. This process is continued until each subset contains only one message. It is easily seen that apart from minor differences (generally in the last digit) this amounts to the same thing as the arithmetic process described above.

10. DISCUSSION

In order to obtain the maximum power transfer from a generator to a load a transformer must in general be introduced so that the generator as seen from the load has the load resistance. The situation here is roughly analogous. The transducer which does the encoding should match the source to the channel in a statistical sense. The source as seen from the channel through the transducer should have the same statistical structure as the source which maximizes the entropy in the channel. The content of Theorem 9 is that, although an exact match is not in general possible, we can approximate it as closely as desired. The ratio of the actual rate of transmission to the capacity C may be called the efficiency of the coding system. This is of course equal to the ratio of the actual entropy of the channel symbols to the maximum possible entropy.

In general, ideal or nearly ideal encoding requires a long delay in the transmitter and receiver. In the noiseless case which we have been considering, the main function of this delay is to allow reasonably good

⁹ Technical Report No. 65, The Research Laboratory of Electronics, M. I. T.

matching of probabilities to corresponding lengths of sequences. With a good code the logarithm of the reciprocal probability of a long message must be proportional to the duration of the corresponding signal, in fact

$$\left| \frac{\log p^{-1}}{T} - C \right|$$

must be small for all but a small fraction of the long messages.

If a source can produce only one particular message its entropy is zero, and no channel is required. For example, a computing machine set up to calculate the successive digits of π produces a definite sequence with no chance element. No channel is required to "transmit" this to another point. One could construct a second machine to compute the same sequence at the point. However, this may be impractical. In such a case we can choose to ignore some or all of the statistical knowledge we have of the source. We might consider the digits of π to be a random sequence in that we construct a system capable of sending any sequence of digits. In a similar way we may choose to use some of our statistical knowledge of English in constructing a code, but not all of it. In such a case we consider the source with the maximum entropy subject to the statistical conditions we wish to retain. The entropy of this source determines the channel capacity which is necessary and sufficient. In the π example the only information retained is that all the digits are chosen from the set 0, 1, . . . , 9. In the case of English one might wish to use the statistical saving possible due to letter frequencies, but nothing else. The maximum entropy source is then the first approximation to English and its entropy determines the required channel capacity.

11. EXAMPLES

As a simple example of some of these results consider a source which produces a sequence of letters chosen from among A, B, C, D with probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}$, successive symbols being chosen independently. We have

$$\begin{aligned} H &= -\left(\frac{1}{2} \log \frac{1}{2} + \frac{1}{4} \log \frac{1}{4} + \frac{2}{8} \log \frac{1}{8}\right) \\ &= \frac{7}{4} \text{ bits per symbol.} \end{aligned}$$

Thus we can approximate a coding system to encode messages from this source into binary digits with an average of $\frac{7}{4}$ binary digit per symbol. In this case we can actually achieve the limiting value by the following code (obtained by the method of the second proof of Theorem 9):

A	0
B	10
C	110
D	111

The average number of binary digits used in encoding a sequence of N symbols will be

$$N(\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{2}{8} \times 3) = \frac{7}{4}N$$

It is easily seen that the binary digits 0, 1 have probabilities $\frac{1}{2}, \frac{1}{2}$ so the H for the coded sequences is one bit per symbol. Since, on the average, we have $\frac{7}{4}$ binary symbols per original letter, the entropies on a time basis are the same. The maximum possible entropy for the original set is $\log 4 = 2$, occurring when A, B, C, D have probabilities $\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}$. Hence the relative entropy is $\frac{7}{8}$. We can translate the binary sequences into the original set of symbols on a two-to-one basis by the following table:

00	A'
01	B'
10	C'
11	D'

This double process then encodes the original message into the same symbols but with an average compression ratio $\frac{7}{8}$.

As a second example consider a source which produces a sequence of A 's and B 's with probability p for A and q for B . If $p \ll q$ we have

$$\begin{aligned} H &= -\log p^p (1-p)^{1-p} \\ &= -p \log p (1-p)^{(1-p)/p} \\ &\doteq p \log \frac{e}{p} \end{aligned}$$

In such a case one can construct a fairly good coding of the message on a 0, 1 channel by sending a special sequence, say 0000, for the infrequent symbol A and then a sequence indicating the *number* of B 's following it. This could be indicated by the binary representation with all numbers containing the special sequence deleted. All numbers up to 16 are represented as usual; 16 is represented by the next binary number after 16 which does not contain four zeros, namely $17 = 10001$, etc.

It can be shown that as $p \rightarrow 0$ the coding approaches ideal provided the length of the special sequence is properly adjusted.

PART II: THE DISCRETE CHANNEL WITH NOISE

11. REPRESENTATION OF A NOISY DISCRETE CHANNEL

We now consider the case where the signal is perturbed by noise during transmission or at one or the other of the terminals. This means that the received signal is not necessarily the same as that sent out by the transmitter. Two cases may be distinguished. If a particular transmitted signal always produces the same received signal, i.e. the received signal is a definite function of the transmitted signal, then the effect may be called distortion. If this function has an inverse—no two transmitted signals producing the same received signal—distortion may be corrected, at least in principle, by merely performing the inverse functional operation on the received signal.

The case of interest here is that in which the signal does not always undergo the same change in transmission. In this case we may assume the received signal E to be a function of the transmitted signal S and a second variable, the noise N .

$$E = f(S, N)$$

The noise is considered to be a chance variable just as the message was above. In general it may be represented by a suitable stochastic process. The most general type of noisy discrete channel we shall consider is a generalization of the finite state noise free channel described previously. We assume a finite number of states and a set of probabilities

$$p_{\alpha, i}(\beta, j).$$

This is the probability, if the channel is in state α and symbol i is transmitted, that symbol j will be received and the channel left in state β . Thus α and β range over the possible states, i over the possible transmitted signals and j over the possible received signals. In the case where successive symbols are independently perturbed by the noise there is only one state, and the channel is described by the set of transition probabilities $p_i(j)$, the probability of transmitted symbol i being received as j .

If a noisy channel is fed by a source there are two statistical processes at work: the source and the noise. Thus there are a number of entropies that can be calculated. First there is the entropy $H(x)$ of the source or of the input to the channel (these will be equal if the transmitter is non-singular). The entropy of the output of the channel, i.e. the received signal, will be denoted by $H(y)$. In the noiseless case $H(y) = H(x)$. The joint entropy of input and output will be $H(xy)$. Finally there are two conditional entropies $H_x(y)$ and $H_y(x)$, the entropy of the output when the input is known and conversely. Among these quantities we have the relations

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x)$$

All of these entropies can be measured on a per-second or a per-symbol basis.

12. EQUIVOCATION AND CHANNEL CAPACITY

If the channel is noisy it is not in general possible to reconstruct the original message or the transmitted signal with *certainty* by any operation on the received signal E . There are, however, ways of transmitting the information which are optimal in combating noise. This is the problem which we now consider.

Suppose there are two possible symbols 0 and 1, and we are transmitting at a rate of 1000 symbols per second with probabilities $p_0 = p_1 = \frac{1}{2}$. Thus our source is producing information at the rate of 1000 bits per second. During transmission the noise introduces errors so that, on the average, 1 in 100 is received incorrectly (a 0 as 1, or 1 as 0). What is the rate of transmission of information? Certainly less than 1000 bits per second since about 1% of the received symbols are incorrect. Our first impulse might be to say the rate is 990 bits per second, merely subtracting the expected number of errors. This is not satisfactory since it fails to take into account the recipient's lack of knowledge of where the errors occur. We may carry it to an extreme case and suppose the noise so great that the received symbols are entirely independent of the transmitted symbols. The probability of receiving 1 is $\frac{1}{2}$ whatever was transmitted and similarly for 0. Then about half of the received symbols are correct due to chance alone, and we would be giving the system credit for transmitting 500 bits per second while actually no information is being transmitted at all. Equally "good" transmission would be obtained by dispensing with the channel entirely and flipping a coin at the receiving point.

Evidently the proper correction to apply to the amount of information transmitted is the amount of this information which is missing in the received signal, or alternatively the uncertainty when we have received a signal of what was actually sent. From our previous discussion of entropy as a measure of uncertainty it seems reasonable to use the conditional entropy of the message, knowing the received signal, as a measure of this missing information. This is indeed the proper definition, as we shall see later. Following this idea the rate of actual transmission, R , would be obtained by subtracting from the rate of production (i.e., the entropy of the source) the average rate of conditional entropy.

$$R = H(x) - H_y(x)$$

The conditional entropy $H_y(x)$ will, for convenience, be called the equivocation. It measures the average ambiguity of the received signal.

In the example considered above, if a 0 is received the *a posteriori* probability that a 0 was transmitted is .99, and that a 1 was transmitted is .01. These figures are reversed if a 1 is received. Hence

$$\begin{aligned} H_y(x) &= - [.99 \log .99 + 0.01 \log 0.01] \\ &= .081 \text{ bits/symbol} \end{aligned}$$

or 81 bits per second. We may say that the system is transmitting at a rate $1000 - 81 = 919$ bits per second. In the extreme case where a 0 is equally likely to be received as a 0 or 1 and similarly for 1, the *a posteriori* probabilities are $\frac{1}{2}$, $\frac{1}{2}$ and

$$\begin{aligned} H_y(x) &= - [\tfrac{1}{2} \log \tfrac{1}{2} + \tfrac{1}{2} \log \tfrac{1}{2}] \\ &= 1 \text{ bit per symbol} \end{aligned}$$

or 1000 bits per second. The rate of transmission is then 0 as it should be.

The following theorem gives a direct intuitive interpretation of the equivocation and also serves to justify it as the unique appropriate measure. We consider a communication system and an observer (or auxiliary device) who can see both what is sent and what is recovered (with errors due to noise). This observer notes the errors in the recovered message and transmits data to the receiving point over a "correction channel" to enable the receiver to correct the errors. The situation is indicated schematically in Fig. 8.

Theorem 10: If the correction channel has a capacity equal to $H_y(x)$ it is possible to so encode the correction data as to send it over this channel and correct all but an arbitrarily small fraction ϵ of the errors. This is not possible if the channel capacity is less than $H_y(x)$.

Roughly then, $H_y(x)$ is the amount of additional information that must be supplied per second at the receiving point to correct the received message.

To prove the first part, consider long sequences of received message M' and corresponding original message M . There will be logarithmically $TH_y(x)$ of the M 's which could reasonably have produced each M' . Thus we have $TH_y(x)$ binary digits to send each T seconds. This can be done with ϵ frequency of errors on a channel of capacity $H_y(x)$.

The second part can be proved by noting, first, that for any discrete chance variables x, y, z

$$H_y(x, z) \geq H_y(x)$$

The left-hand side can be expanded to give

$$\begin{aligned} H_y(z) + H_{yz}(x) &\geq H_y(x) \\ H_{yz}(x) &\geq H_y(x) - H_y(z) \geq H_y(x) - H(z) \end{aligned}$$

If we identify x as the output of the source, y as the received signal and z as the signal sent over the correction channel, then the right-hand side is the equivocation less the rate of transmission over the correction channel. If the capacity of this channel is less than the equivocation the right-hand side will be greater than zero and $H_{yz}(x) \geq 0$. But this is the uncertainty of what was sent, knowing both the received signal and the correction signal. If this is greater than zero the frequency of errors cannot be arbitrarily small.

Example:

Suppose the errors occur at random in a sequence of binary digits: probability p that a digit is wrong and $q = 1 - p$ that it is right. These errors can be corrected if their position is known. Thus the correction channel need only send information as to these positions. This amounts to trans-

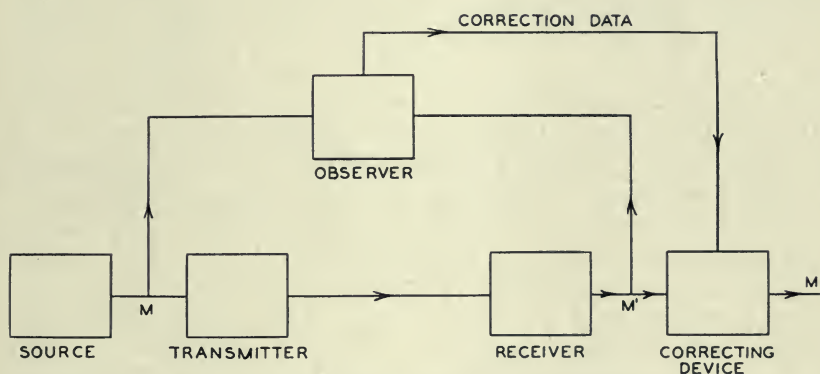


Fig. 8—Schematic diagram of a correction system.

mitting from a source which produces binary digits with probability p for 1 (correct) and q for 0 (incorrect). This requires a channel of capacity

$$-[p \log p + q \log q]$$

which is the equivocation of the original system.

The rate of transmission R can be written in two other forms due to the identities noted above. We have

$$\begin{aligned} R &= H(x) - H_y(x) \\ &= H(y) - H_x(y) \\ &= H(x) + H(y) - H(x, y). \end{aligned}$$

The first defining expression has already been interpreted as the amount of information sent less the uncertainty of what was sent. The second meas-

ures the amount received less the part of this which is due to noise. The third is the sum of the two amounts less the joint entropy and therefore in a sense is the number of bits per second common to the two. Thus all three expressions have a certain intuitive significance.

The capacity C of a noisy channel should be the maximum possible rate of transmission, i.e., the rate when the source is properly matched to the channel. We therefore define the channel capacity by

$$C = \text{Max } (H(x) - H_y(x))$$

where the maximum is with respect to all possible information sources used as input to the channel. If the channel is noiseless, $H_y(x) = 0$. The definition is then equivalent to that already given for a noiseless channel since the maximum entropy for the channel is its capacity.

13. THE FUNDAMENTAL THEOREM FOR A DISCRETE CHANNEL WITH NOISE

It may seem surprising that we should define a definite capacity C for a noisy channel since we can never send certain information in such a case. It is clear, however, that by sending the information in a redundant form the probability of errors can be reduced. For example, by repeating the message many times and by a statistical study of the different received versions of the message the probability of errors could be made very small. One would expect, however, that to make this probability of errors approach zero, the redundancy of the encoding must increase indefinitely, and the rate of transmission therefore approach zero. This is by no means true. If it were, there would not be a very well defined capacity, but only a capacity for a given frequency of errors, or a given equivocation; the capacity going down as the error requirements are made more stringent. Actually the capacity C defined above has a very definite significance. It is possible to send information at the rate C through the channel *with as small a frequency of errors or equivocation as desired* by proper encoding. This statement is not true for any rate greater than C . If an attempt is made to transmit at a higher rate than C , say $C + R_1$, then there will necessarily be an equivocation equal to a greater than the excess R_1 . Nature takes payment by requiring just that much uncertainty, so that we are not actually getting any more than C through correctly.

The situation is indicated in Fig. 9. The rate of information into the channel is plotted horizontally and the equivocation vertically. Any point above the heavy line in the shaded region can be attained and those below cannot. The points on the line cannot in general be attained, but there will usually be two points on the line that can.

These results are the main justification for the definition of C and will now be proved.

Theorem 11. Let a discrete channel have the capacity C and a discrete source the entropy per second H . If $H \leq C$ there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors (or an arbitrarily small equivocation). If $H > C$ it is possible to encode the source so that the equivocation is less than $H - C + \epsilon$ where ϵ is arbitrarily small. There is no method of encoding which gives an equivocation less than $H - C$.

The method of proving the first part of this theorem is not by exhibiting a coding method having the desired properties, but by showing that such a code must exist in a certain group of codes. In fact we will average the frequency of errors over this group and show that this average can be made less than ϵ . If the average of a set of numbers is less than ϵ there must exist at least one in the set which is less than ϵ . This will establish the desired result.

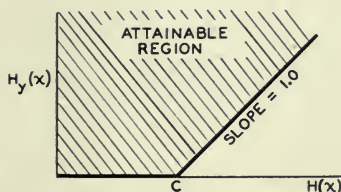


Fig. 9—The equivocation possible for a given input entropy to a channel.

The capacity C of a noisy channel has been defined as

$$C = \text{Max } (H(x) - H_y(x))$$

where x is the input and y the output. The maximization is over all sources which might be used as input to the channel.

Let S_0 be a source which achieves the maximum capacity C . If this maximum is not actually achieved by any source let S_0 be a source which approximates to giving the maximum rate. Suppose S_0 is used as input to the channel. We consider the possible transmitted and received sequences of a long duration T . The following will be true:

1. The transmitted sequences fall into two classes, a high probability group with about $2^{TH(x)}$ members and the remaining sequences of small total probability.
2. Similarly the received sequences have a high probability set of about $2^{TH(y)}$ members and a low probability set of remaining sequences.
3. Each high probability output could be produced by about $2^{TH_y(x)}$ inputs. The probability of all other cases has a small total probability.

All the ϵ 's and δ 's implied by the words "small" and "about" in these statements approach zero as we allow T to increase and S_0 to approach the maximizing source.

The situation is summarized in Fig. 10 where the input sequences are points on the left and output sequences points on the right. The fan of cross lines represents the range of possible causes for a typical output.

Now suppose we have another source producing information at rate R with $R < C$. In the period T this source will have 2^{TR} high probability outputs. We wish to associate these with a selection of the possible channel

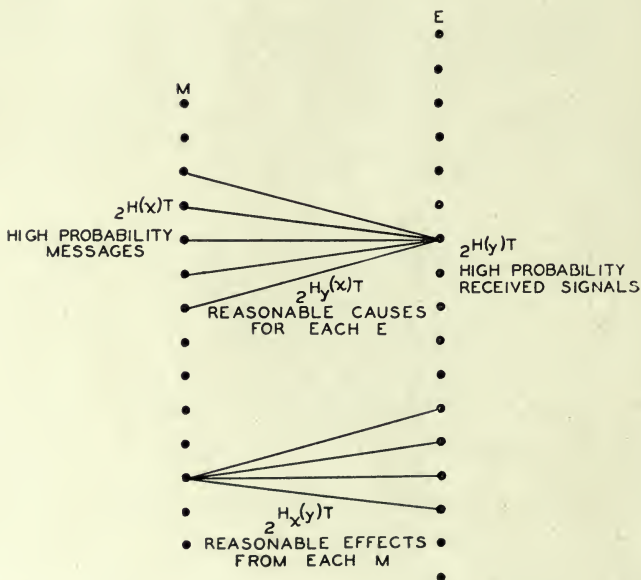


Fig. 10—Schematic representation of the relations between inputs and outputs in a channel.

inputs in such a way as to get a small frequency of errors. We will set up this association in all possible ways (using, however, only the high probability group of inputs as determined by the source S_0) and average the frequency of errors for this large class of possible coding systems. This is the same as calculating the frequency of errors for a random association of the messages and channel inputs of duration T . Suppose a particular output y_1 is observed. What is the probability of more than one message in the set of possible causes of y_1 ? There are 2^{TR} messages distributed at random in $2^{TH(x)}$ points. The probability of a particular point being a message is thus

$$2^{T(R-H(x))}$$

The probability that none of the points in the fan is a message (apart from the actual originating message) is

$$P = [1 - 2^{T(R-H(x))}]^{2^{TH_y(x)}}$$

Now $R < H(x) - H_y(x)$ so $R - H(x) = -H_y(x) - \eta$ with η positive. Consequently

$$P = [1 - 2^{-TH_y(x) - T\eta}]^{2^{TH_y(x)}}$$

approaches (as $T \rightarrow \infty$)

$$1 - 2^{-T\eta}.$$

Hence the probability of an error approaches zero and the first part of the theorem is proved.

The second part of the theorem is easily shown by noting that we could merely send C bits per second from the source, completely neglecting the remainder of the information generated. At the receiver the neglected part gives an equivocation $H(x) - C$ and the part transmitted need only add ϵ . This limit can also be attained in many other ways, as will be shown when we consider the continuous case.

The last statement of the theorem is a simple consequence of our definition of C . Suppose we can encode a source with $R = C + a$ in such a way as to obtain an equivocation $H_y(x) = a - \epsilon$ with ϵ positive. Then $R = H(x) = C + a$ and

$$H(x) - H_y(x) = C + \epsilon$$

with ϵ positive. This contradicts the definition of C as the maximum of $H(x) - H_y(x)$.

Actually more has been proved than was stated in the theorem. If the average of a set of numbers is within ϵ of their maximum, a fraction of at most $\sqrt{\epsilon}$ can be more than $\sqrt{\epsilon}$ below the maximum. Since ϵ is arbitrarily small we can say that almost all the systems are arbitrarily close to the ideal.

14. DISCUSSION

The demonstration of theorem 11, while not a pure existence proof, has some of the deficiencies of such proofs. An attempt to obtain a good approximation to ideal coding by following the method of the proof is generally impractical. In fact, apart from some rather trivial cases and certain limiting situations, no explicit description of a series of approximation to the ideal has been found. Probably this is no accident but is related to the difficulty of giving an explicit construction for a good approximation to a random sequence.

An approximation to the ideal would have the property that if the signal is altered in a reasonable way by the noise, the original can still be recovered. In other words the alteration will not in general bring it closer to another reasonable signal than the original. This is accomplished at the cost of a certain amount of redundancy in the coding. The redundancy must be introduced in the proper way to combat the particular noise structure involved. However, any redundancy in the source will usually help if it is utilized at the receiving point. In particular, if the source already has a certain redundancy and no attempt is made to eliminate it in matching to the channel, this redundancy will help combat noise. For example, in a noiseless telegraph channel one could save about 50% in time by proper encoding of the messages. This is not done and most of the redundancy of English remains in the channel symbols. This has the advantage, however, of allowing considerable noise in the channel. A sizable fraction of the letters can be received incorrectly and still reconstructed by the context. In fact this is probably not a bad approximation to the ideal in many cases, since the statistical structure of English is rather involved and the reasonable English sequences are not too far (in the sense required for theorem) from a random selection.

As in the noiseless case a delay is generally required to approach the ideal encoding. It now has the additional function of allowing a large sample of noise to affect the signal before any judgment is made at the receiving point as to the original message. Increasing the sample size always sharpens the possible statistical assertions.

The content of theorem 11 and its proof can be formulated in a somewhat different way which exhibits the connection with the noiseless case more clearly. Consider the possible signals of duration T and suppose a subset of them is selected to be used. Let those in the subset all be used with equal probability, and suppose the receiver is constructed to select, as the original signal, the most probable cause from the subset, when a perturbed signal is received. We define $N(T, q)$ to be the maximum number of signals we can choose for the subset such that the probability of an incorrect interpretation is less than or equal to q .

Theorem 12: $\lim_{T \rightarrow \infty} \frac{\log N(T, q)}{T} = C$, where C is the channel capacity, provided that q does not equal 0 or 1.

In other words, no matter how we set our limits of reliability, we can distinguish reliably in time T enough messages to correspond to about CT bits, when T is sufficiently large. Theorem 12 can be compared with the definition of the capacity of a noiseless channel given in section 1.

15. EXAMPLE OF A DISCRETE CHANNEL AND ITS CAPACITY

A simple example of a discrete channel is indicated in Fig. 11. There are three possible symbols. The first is never affected by noise. The second and third each have probability p of coming through undisturbed, and q of being changed into the other of the pair. We have (letting $\alpha = -[p \log$

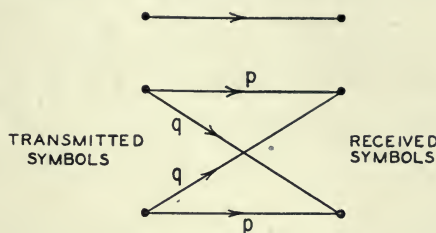


Fig. 11—Example of a discrete channel.

$p + q \log q]$ and P and Q be the probabilities of using the first or second symbols)

$$H(x) = -P \log P - 2Q \log Q$$

$$H_y(x) = 2Q\alpha$$

We wish to choose P and Q in such a way as to maximize $H(x) - H_y(x)$, subject to the constraint $P + 2Q = 1$. Hence we consider

$$U = -P \log P - 2Q \log Q - 2Q\alpha + \lambda(P + 2Q)$$

$$\frac{\partial U}{\partial P} = -1 - \log P + \lambda = 0$$

$$\frac{\partial U}{\partial Q} = -2 - 2 \log Q - 2\alpha + 2\lambda = 0.$$

Eliminating λ

$$\log P = \log Q + \alpha$$

$$P = Qe^\alpha = Q\beta$$

$$P = \frac{\beta}{\beta + 2} \quad Q = \frac{1}{\beta + 2}.$$

The channel capacity is then

$$C = \log \frac{\beta + 2}{\beta}.$$

Note how this checks the obvious values in the cases $p = 1$ and $p = \frac{1}{2}$. In the first, $\beta = 1$ and $C = \log 3$, which is correct since the channel is then noiseless with three possible symbols. If $p = \frac{1}{2}$, $\beta = 2$ and $C = \log 2$. Here the second and third symbols cannot be distinguished at all and act together like one symbol. The first symbol is used with probability $P = \frac{1}{2}$ and the second and third together with probability $\frac{1}{2}$. This may be distributed in any desired way and still achieve the maximum capacity.

For intermediate values of p the channel capacity will lie between $\log 2$ and $\log 3$. The distinction between the second and third symbols conveys some information but not as much as in the noiseless case. The first symbol is used somewhat more frequently than the other two because of its freedom from noise.

16. THE CHANNEL CAPACITY IN CERTAIN SPECIAL CASES

If the noise affects successive channel symbols independently it can be described by a set of transition probabilities p_{ij} . This is the probability, if symbol i is sent, that j will be received. The maximum channel rate is then given by the maximum of

$$\sum_{i,j} P_i p_{ij} \log \sum_i P_i p_{ij} - \sum_{i,j} P_i p_{ij} \log p_{ij}$$

where we vary the P_i subject to $\sum P_i = 1$. This leads by the method of Lagrange to the equations,

$$\sum_j p_{sj} \log \frac{p_{sj}}{\sum_i P_i p_{ij}} = \mu \quad s = 1, 2, \dots$$

Multiplying by P_s and summing on s shows that $\mu = -C$. Let the inverse of p_{sj} (if it exists) be h_{st} so that $\sum_s h_{st} p_{sj} = \delta_{tj}$. Then:

$$\sum_{s,j} h_{st} p_{sj} \log p_{sj} - \log \sum_i P_i p_{it} = -C \sum_s h_{st}$$

Hence:

$$\sum_i P_i p_{it} = \exp [C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{sj}]$$

or,

$$P_i = \sum_t h_{it} \exp [C \sum_s h_{st} + \sum_{s,j} h_{st} p_{sj} \log p_{sj}].$$

This is the system of equations for determining the maximizing values of P_i , with C to be determined so that $\sum P_i = 1$. When this is done C will be the channel capacity, and the P_i the proper probabilities for the channel symbols to achieve this capacity.

If each input symbol has the same set of probabilities on the lines emerging from it, and the same is true of each output symbol, the capacity can be easily calculated. Examples are shown in Fig. 12. In such a case $H_x(y)$ is independent of the distribution of probabilities on the input symbols, and is given by $-\sum p_i \log p_i$ where the p_i are the values of the transition probabilities from any input symbol. The channel capacity is

$$\begin{aligned} \text{Max } [H(y) - H_x(y)] \\ = \text{Max } H(y) + \sum p_i \log p_i. \end{aligned}$$

The maximum of $H(y)$ is clearly $\log m$ where m is the number of output

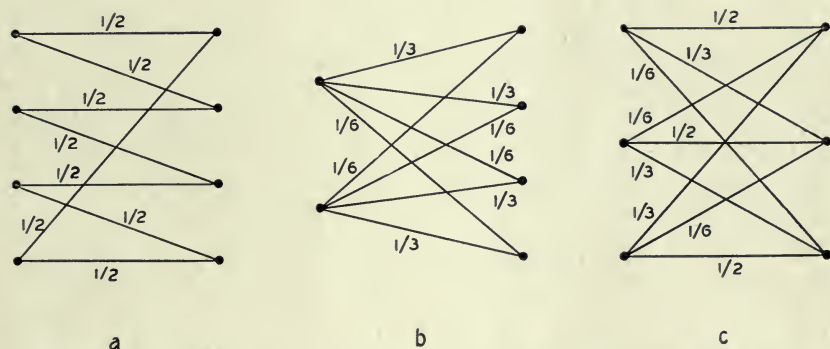


Fig. 12—Examples of discrete channels with the same transition probabilities for each input and for each output.

symbols, since it is possible to make them all equally probable by making the input symbols equally probable. The channel capacity is therefore

$$C = \log m + \sum p_i \log p_i.$$

In Fig. 12a it would be

$$C = \log 4 - \log 2 = \log 2.$$

This could be achieved by using only the 1st and 3d symbols. In Fig. 12b

$$\begin{aligned} C &= \log 4 - \frac{2}{3} \log 3 - \frac{1}{3} \log 6 \\ &= \log 4 - \log 3 - \frac{1}{3} \log 2 \\ &= \log \frac{1}{3} 2^{\frac{4}{3}}. \end{aligned}$$

In Fig. 12c we have

$$\begin{aligned} C &= \log 3 - \frac{1}{2} \log 2 - \frac{1}{3} \log 3 - \frac{1}{6} \log 6 \\ &= \log \frac{3}{2^{\frac{1}{2}} 3^{\frac{1}{3}} 6^{\frac{1}{6}}}. \end{aligned}$$

Suppose the symbols fall into several groups such that the noise never causes a symbol in one group to be mistaken for a symbol in another group. Let the capacity for the n th group be C_n when we use only the symbols in this group. Then it is easily shown that, for best use of the entire set, the total probability P_n of all symbols in the n th group should be

$$P_n = \frac{2^{C_n}}{\sum 2^{C_n}}.$$

Within a group the probability is distributed just as it would be if these were the only symbols being used. The channel capacity is

$$C = \log \sum 2^{C_n}.$$

17. AN EXAMPLE OF EFFICIENT CODING

The following example, although somewhat unrealistic, is a case in which exact matching to a noisy channel is possible. There are two channel symbols, 0 and 1, and the noise affects them in blocks of seven symbols. A block of seven is either transmitted without error, or exactly one symbol of the seven is incorrect. These eight possibilities are equally likely. We have

$$\begin{aligned} C &= \text{Max} [H(y) - H_x(y)] \\ &= \frac{1}{7} [7 + \frac{8}{8} \log \frac{1}{8}] \\ &= \frac{4}{7} \text{ bits/symbol.} \end{aligned}$$

An efficient code, allowing complete correction of errors and transmitting at the rate C , is the following (found by a method due to R. Hamming):

Let a block of seven symbols be $X_1, X_2, \dots, X_7$. Of these X_3, X_5, X_6 and X_7 are message symbols and chosen arbitrarily by the source. The other three are redundant and calculated as follows:

X_4 is chosen to make $\alpha = X_4 + X_5 + X_6 + X_7$ even

X_2 " " " " $\beta = X_2 + X_3 + X_5 + X_7$ "

X_1 " " " " $\gamma = X_1 + X_3 + X_5 + X_7$ "

When a block of seven is received α, β and γ are calculated and if even called zero, if odd called one. The binary number $\alpha \beta \gamma$ then gives the subscript of the X_i that is incorrect (if 0 there was no error).

APPENDIX 1

THE GROWTH OF THE NUMBER OF BLOCKS OF SYMBOLS WITH A FINITE STATE CONDITION

Let $N_i(L)$ be the number of blocks of symbols of length L ending in state i . Then we have

$$N_j(L) = \sum_{i \neq s} N_i(L - b_{ij}^{(s)})$$

where $b_{ij}^1, b_{ij}^2, \dots, b_{ij}^m$ are the length of the symbols which may be chosen in state i and lead to state j . These are linear difference equations and the behavior as $L \rightarrow \infty$ must be of the type

$$N_j = A_j W^L$$

Substituting in the difference equation

$$A_j W^L = \sum_{i,s} A_i W^{L-b_{ij}^{(s)}}$$

or

$$A_j = \sum_{i,s} A_i W^{-b_{ij}^{(s)}}$$

$$\sum_i \left(\sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right) A_i = 0.$$

For this to be possible the determinant

$$D(W) = |a_{ij}| = \left| \sum_s W^{-b_{ij}^{(s)}} - \delta_{ij} \right|$$

must vanish and this determines W , which is, of course, the largest real root of $D = 0$.

The quantity C is then given by

$$C = \lim_{L \rightarrow \infty} \frac{\log \sum A_j W^L}{L} = \log W$$

and we also note that the same growth properties result if we require that all blocks start in the same (arbitrarily chosen) state.

APPENDIX 2

$$\text{DERIVATION OF } H = -\sum p_i \log p_i$$

Let $H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) = A(n)$. From condition (3) we can decompose a choice from s^m equally likely possibilities into a series of m choices each from s equally likely possibilities and obtain

$$A(s^m) = m A(s)$$

Similarly

$$A(l^n) = n A(l)$$

We can choose n arbitrarily large and find an m to satisfy

$$s^m \leq l^n < s^{(m+1)}$$

Thus, taking logarithms and dividing by $n \log s$,

$$\frac{m}{n} \leq \frac{\log t}{\log s} \leq \frac{m}{n} + \frac{1}{n} \quad \text{or} \quad \left| \frac{m}{n} - \frac{\log t}{\log s} \right| < \epsilon$$

where ϵ is arbitrarily small.

Now from the monotonic property of $A(n)$

$$A(s^m) \leq A(t^n) \leq A(s^{m+1})$$

$$m A(s) \leq n A(t) \leq (m+1) A(s)$$

Hence, dividing by $n A(s)$,

$$\frac{m}{n} \leq \frac{A(t)}{A(s)} \leq \frac{m}{n} + \frac{1}{n} \quad \text{or} \quad \left| \frac{m}{n} - \frac{A(t)}{A(s)} \right| < \epsilon$$

$$\left| \frac{A(t)}{A(s)} - \frac{\log t}{\log s} \right| \leq 2\epsilon \quad A(t) = -K \log t$$

where K must be positive to satisfy (2).

Now suppose we have a choice from n possibilities with commensurable probabilities $p_i = \frac{n_i}{\Sigma n_i}$ where the n_i are integers. We can break down a choice from Σn_i possibilities into a choice from n possibilities with probabilities $p_i \dots p_n$ and then, if the i th was chosen, a choice from n_i with equal probabilities. Using condition 3 again, we equate the total choice from Σn_i as computed by two methods

$$K \log \Sigma n_i = H(p_1, \dots, p_n) + K \Sigma p_i \log n_i$$

Hence

$$\begin{aligned} H &= K [\Sigma p_i \log \Sigma n_i - \Sigma p_i \log n_i] \\ &= -K \Sigma p_i \log \frac{n_i}{\Sigma n_i} = -K \Sigma p_i \log p_i. \end{aligned}$$

If the p_i are incommensurable, they may be approximated by rationals and the same expression must hold by our continuity assumption. Thus the expression holds in general. The choice of coefficient K is a matter of convenience and amounts to the choice of a unit of measure.

APPENDIX 3

THEOREMS ON ERGODIC SOURCES

If it is possible to go from any state with $P > 0$ to any other along a path of probability $p > 0$, the system is ergodic and the strong law of large numbers can be applied. Thus the number of times a given path p_{ij} in the net-

work is traversed in a long sequence of length N is about proportional to the probability of being at i and then choosing this path, $P_i p_{ij} N$. If N is large enough the probability of percentage error $\pm \delta$ in this is less than ϵ so that for all but a set of small probability the actual numbers lie within the limits

$$(P_i p_{ij} \pm \delta) N$$

Hence nearly all sequences have a probability p given by

$$p = \prod p_{ij}^{(P_i p_{ij} \pm \delta) N}$$

and $\frac{\log p}{N}$ is limited by

$$\frac{\log p}{N} = \Sigma (P_i p_{ij} \pm \delta) \log p_{ij}$$

or

$$\left| \frac{\log p}{N} - \Sigma P_i p_{ij} \log p_{ij} \right| < \eta.$$

This proves theorem 3.

Theorem 4 follows immediately from this on calculating upper and lower bounds for $n(q)$ based on the possible range of values of p in Theorem 3.

In the mixed (not ergodic) case if

$$L = \Sigma p_i L_i$$

and the entropies of the components are $H_1 \geq H_2 \geq \dots \geq H_n$ we have the

Theorem: $\lim_{N \rightarrow \infty} \frac{\log n(q)}{N} = \varphi(q)$ is a decreasing step function,

$$\varphi(q) = H_s \quad \text{in the interval} \quad \sum_1^{s-1} \alpha_i < q < \sum_1^s \alpha_i.$$

To prove theorems 5 and 6 first note that F_N is monotonic decreasing because increasing N adds a subscript to a conditional entropy. A simple substitution for $p_{Bi}(S_j)$ in the definition of F_N shows that

$$F_N = N G_N - (N - 1) G_{N-1}$$

and summing this for all N gives $G_N = \frac{1}{N} \Sigma F_N$. Hence $G_N \geq F_N$ and G_N monotonic decreasing. Also they must approach the same limit. By using theorem 3 we see that $\lim_{N \rightarrow \infty} G_N = H$.

APPENDIX 4

MAXIMIZING THE RATE FOR A SYSTEM OF CONSTRAINTS

Suppose we have a set of constraints on sequences of symbols that is of the finite state type and can be represented therefore by a linear graph.

Let $\ell_{ij}^{(s)}$ be the lengths of the various symbols that can occur in passing from state i to state j . What distribution of probabilities P_i for the different states and $p_{ij}^{(s)}$ for choosing symbol s in state i and going to state j maximizes the rate of generating information under these constraints? The constraints define a discrete channel and the maximum rate must be less than or equal to the capacity C of this channel, since if all blocks of large length were equally likely, this rate would result, and if possible this would be best. We will show that this rate can be achieved by proper choice of the P_i and $p_{ij}^{(s)}$.

The rate in question is

$$\frac{-\sum P_i p_{ij}^{(s)} \log p_{ij}^{(s)}}{\sum P_i p_{ij}^{(s)} \ell_{ij}^{(s)}} = \frac{N}{M}.$$

Let $\ell_{ij} = \sum_s \ell_{ij}^{(s)}$. Evidently for a maximum $p_{ij}^{(s)} = k \exp \ell_{ij}^{(s)}$. The constraints on maximization are $\sum P_i = 1$, $\sum_j p_{ij} = 1$, $\sum P_i (p_{ij} - \delta_{ij}) = 0$.

Hence we maximize

$$U = \frac{-\sum P_i p_{ij} \log p_{ij}}{\sum P_i p_{ij} \ell_{ij}} + \lambda \sum_i P_i + \sum \mu_i p_{ij} + \sum \eta_j P_i (p_{ij} - \delta_{ij})$$

$$\frac{\partial U}{\partial p_{ij}} = -\frac{MP_i(1 + \log p_{ij}) + NP_i \ell_{ij}}{M^2} + \lambda + \mu_i + \eta_j P_i = 0.$$

Solving for p_{ij}

$$p_{ij} = A_i B_j D^{-\ell_{ij}}.$$

Since

$$\sum_j p_{ij} = 1, \quad A_i^{-1} = \sum_j B_j D^{-\ell_{ij}}$$

$$p_{ij} = \frac{B_j D^{-\ell_{ij}}}{\sum_s B_s D^{-\ell_{is}}}.$$

The correct value of D is the capacity C and the B_j are solutions of

$$B_i = \sum B_j C^{-\ell_{ij}}$$

for then

$$p_{ij} = \frac{B_j}{B_i} C^{-\ell_{ij}}$$

$$\sum P_i \frac{B_j}{B_i} C^{-\ell_{ij}} = P_j$$

or

$$\sum \frac{P_i}{B_i} C^{-\ell_{ij}} = \frac{P_j}{B_j}$$

So that if λ_i satisfy

$$\sum \gamma_i C^{-\ell_{ij}} = \gamma_j$$

$$P_i = B_i \gamma_i$$

Both of the sets of equations for B_i and γ_i can be satisfied since C is such that

$$|C^{-\ell_{ij}} - \delta_{ij}| = 0$$

In this case the rate is

$$\begin{aligned} & - \frac{\sum P_i p_{ij} \log \frac{B_j}{B_i} C^{-\ell_{ij}}}{\sum P_i p_{ij} \ell_{ij}} \\ & = C - \frac{\sum P_i p_{ij} \log \frac{B_j}{B_i}}{\sum P_i p_{ij} \ell_{ij}} \end{aligned}$$

but

$$\sum P_i p_{ij} (\log B_j - \log B_i) = \sum_j P_j \log B_j - \sum_i P_i \log B_i = 0$$

Hence the rate is C and as this could never be exceeded this is the maximum, justifying the assumed solution.

(To be continued)

An Aspect of the Dialing Behavior of Subscribers and Its Effect on the Trunk Plant

By CHARLES CLOS

INTRODUCTION

DURING the war it became necessary for the Bell System Companies to lower many service standards. Among these was the standard for the provision of trunks for handling subscriber-dialed calls. In the interest of economy the number of trunks for a given volume of traffic was lowered. It is evident that for any given case there is a lower limit to the number of trunks that should be provided for handling subscriber-dialed calls. Below this limit congestion of calls gets beyond control. The control of congestion is important. In the case of operator-handled calls it is possible to control congestion by filing tickets and placing calls in an orderly fashion. In the case of subscriber-dialed calls the subscriber may with impunity make many, indeed very many, successive dialing attempts to complete a call that is blocked due to a shortage of trunks. If, in a particular office enough subscribers do this simultaneously, a sender shortage may develop with its resulting reaction on the whole office.

From the foregoing it is evident that the standard of service for providing trunks in trunk groups handling subscriber-dialed calls is of importance. During the war years, the New York Telephone Company undertook a study to determine the limits below which it would be undesirable to degrade the service. This study was designed to test the reasonableness of the reduction in the inter-office trunk standard from the pre-war basis of providing enough trunks to delay only one out of a hundred calls in the busy hour to a wartime basis of providing enough trunks to delay two calls in every hundred during the busy hour. The conclusion from this study was that it was safe to use wartime standards.

The study reported herein is an analysis of the effect of repeated attempts when subscriber-dialed calls are blocked due to trunk shortages. The data upon which the results are based indicate that dial subscribers after encountering a *busy* condition make new attempts sooner and much more often than has been generally believed. The results indicate that one can reconstruct what happens when trunk groups carrying subscriber-dialed calls encounter serious overloads and that trunk capacity tables for such situations can be developed.

The study is based on extensive service observations taken at the New

York City Service Observing Bureaus during the winter of 1943-44. These observations dealt with the behavior of subscribers who encounter a *busy* on a dialed call. This behavior is assumed to apply to the situation when subscribers encounter an *all-trunks-busy* condition.

INADEQUACY OF THE POISSON AND ERLANG B FORMULAE TO EXPRESS
THE SITUATION WHEN SHORTAGES OCCUR IN TRUNK GROUPS
HANDLING SUBSCRIBER DIALED CALLS

In connection with the provision of trunks in the exchange plant, two sets of trunk-call-carrying-capacity tables are currently in use. One set of these tables is computed from the Poisson Formula and the other from the Erlang B Formula. The Poisson tables are used for trunk groups carrying non-alternate route traffic, whereas the Erlang B tables are used for trunk groups carrying traffic subject to alternate routing.

The assumption underlying the Poisson Formula, when a shortage of trunks occurs, is that of a partial delay. A call which encounters *all trunks busy* waits but not longer than a holding time interval for a trunk to become available.

The corresponding assumption underlying the Erlang B Formula is that of no delay. A call which encounters *all trunks busy* is cleared out. The call may be abandoned by the subscriber or advanced to an alternate route.

With respect to non-alternate route trunk groups handling subscriber dialed calls neither of the above two assumptions is realized in practice. When all trunks are busy, the dial equipment is arranged to return an *all-trunks-busy* signal to the subscriber rather than hold the call pending the outcome of a subsequent test for an idle trunk. The subscriber upon encountering an *all-trunks-busy* signal does not necessarily abandon the call. In most cases he redials the call.

The degree by which the assumptions are not realized depends upon the relative number of trunks that are provided for a given volume of traffic. For instance if, during an hour, 150 calls having an average holding time of 100 seconds are submitted to ten trunks and an equivalent volume of traffic is submitted to five trunks, the following theoretical results follow from the Poisson and Erlang B Formulae:—

TABLE I
THEORETICAL RESULTS FROM POISSON AND ERLANG B FORMULAE

150 Calls of 100 Seconds Average Holding Time Submitted during an hour to	Number of Calls that Are Delayed on the Basis of the Poisson Formula	Number of Calls that Are Cleared Out on the Basis of the Erlang B Formula
10 trunks	1.6	1.0
5 trunks	60.6	32.0

The values in Table I indicate that, when a liberal number of trunks, i.e., ten trunks, is provided, the numerical difference between the results of the two formulae is small and the results of either formula can be used as an approximation of the number of calls affected by an *all-trunks-busy* condition. There are undoubtedly repetitious attempts, but because the number is small their effect can be neglected.

When, however, there is a serious shortage of trunks, as when only five trunks are provided, the numerical difference between the theoretical results of the two formulae is large. In addition, the repetitious attempts will be too numerous to ignore. Some of the repetitious attempts will encounter *all trunks busy* again and again. Other repetitious attempts will seize idle trunks thereby causing new calls to encounter *all trunks busy*. The effect is cumulative. Neither the Poisson nor the Erlang B Formula indicates to what extent the repetitious attempts take place nor their effect. A preliminary glimpse at the results of this study indicates that 150 calls of 100 seconds average holding time when submitted during an hour to five trunks become inflated by 99 repetitious attempts and appear as 249 calls being submitted to the trunks. Of these 249 calls, 108 encounter *all trunks busy*. Of the 108 calls, 99 become the aforementioned repetitious attempts and nine are abandoned. It is evident that neither formula presents this picture. For studies considering the effect of overloads due to trunk shortages, this is the type of information needed. A new approach is required to obtain such data. To do this, it is desirable to examine the habits of dial subscribers who have encountered *busies*.

THE DIALING BEHAVIOR OF SUBSCRIBERS UPON ENCOUNTERING A *BUSY*

In order to investigate the grade of service given to dial subscribers when trunk shortages occur it is desirable to know something about their behavior when they encounter *all-trunks-busy* signals. Specifically there are four items that need investigation; these are:—

1. How soon after encountering an *all-trunks-busy* signal does the subscriber redial his call?
2. What percentage of the subscribers make subsequent attempts?
3. How do the time intervals between successive subsequent attempts compare with each other; that is, are they about the same or do they differ widely?
4. What differences, if any, exist between classes of subscribers?

The first three items are answered from the results of service observations. The fourth item is answered indirectly.

The service observations consisted of 1,107 cases where line *busies* were observed (except for 35 cases of *all-trunks-busy* signals). Observations on

line *busies* were used instead of *all-trunks-busy* signals because it would have taken too long to obtain sufficient observations, because it is undesirable to artificially degrade the service in order to obtain sufficient observations and because it is assumed that the average subscriber does not recognize the difference between a *busy* and overflow signal. It is considered that the data, while collected for *busy* signals, accurately represent the situation with regard to overflow signals.

Beginning on December 22, 1943 and ending on February 29, 1944, a special record of 1,107 subscriber-dialed calls, where line *busies* were observed, was taken at the three New York City service observation bureaus. Up to a point, regular service observation practices were followed and the regular service observing data concerning the calls were entered on the service-observing records. The data concerning the line *busies* were entered on a special form. This form is shown below. Instructions for the observers accompanied these forms; these instructions follow the form.

Form S.O. 171

SPECIAL RECORDS—BUSY CALLS

Calling No.

Date.

Enter in space under attempt number, the cumulative seconds from the start of the original attempt to the start of the attempt indicated. In addition for the last attempt show disposition.

Attempt Number

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20

Disposition of the call.

Data for attempts over 20 should be entered on the reverse side.

Special service observing form used to collect data concerning the dialing behavior of subscribers upon encountering a *busy*.

Instructions Applying to the Use of Form S.O. 171

These instructions apply to the use of Form S.O. 171 which has been developed in connection with a study of the behavior of customers upon encountering a *busy* signal.

This study will not include observations originating on P.B.X. trunks or on coin lines. On all other calls encountering a *busy* signal or an overflow signal the observer will hold the line in the observing position until one of the following conditions occurs:

- (1) Call is disposed of by reaching the desired number.—Code OK

- (2) 10 minutes have elapsed since the last attempt for the desired number.—Code AB
- (3) Call is disposed of by being given to the operator.—Code PR
- (4) Call is disposed of by receiving a "Don't answer" on an attempt to reach the desired number—Code DA

All attempts made during the period that an observation is ordinarily held will be entered on the service observing detail sheets in the regular way. In addition, these entries and entries showing any other attempts to reach the desired number together with the proper code listed above to show the final disposition of the call will be recorded on Form S.O. 171.

In order to minimize the number of cases not completed at the end of an observer's trick, no cases will be recorded on the special record on which the original *busy* signal is received after $\frac{1}{2}$ hour prior to the finish of any trick.

From the instructions it may be noted that observations originating on P.B.X. trunks or on coin lines were not included. The reason for this is, when a *busy* is observed on a call originating on a P.B.X. trunk the subsequent attempt might be made on one of the other P.B.X. trunks, thus the subsequent attempt would be missed. Also, at a P.B.X. two extensions may place calls, within a few seconds of each other, to the same *busy* line. The service observations on any one trunk might therefore be a mixture of attempts involving two or more calls. When a *busy* is observed on a call made from a coin line, the calling party will in many instances vacate the coin box in favor of someone else, and the subsequent attempt may then be made from another coin line. For these reasons the observations were restricted to business and residential individual lines and to two-party lines (12 observations were on two-party lines).

It may also be noted that the observers were instructed to hold the line in the observing position until ten minutes have elapsed since the last attempt for the desired number. This was a departure from regular service observing practices when a line is held until 1 minute has elapsed.

Table II is a tabulation of the data observed at the Manhattan Service Observing Bureau on Manhattan dial subscriber lines. The observations are arranged in the order of increasing magnitude of the time intervals between the start of the first attempt and the start of the second attempt. Of interest is observation number 197 where a subscriber made 25 attempts in about an hour.

Data similar to that observed on Manhattan dial subscriber lines were likewise observed on Bronx-Westchester and on Brooklyn-Queens dial subscriber lines.

Figure 1(a) shows graphically the data listed in Table II. This graph shows, by dots, the cumulative percentage of the 451 Manhattan observa-

TABLE II
RESULTS OF OBSERVATIONS ON 451 DIAL SUBSCRIBERS IN MANHATTAN

Seconds elapsing between start of previous attempt and start of attempt listed below:

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
1	0	13	24	13	11	20					81	O.K.
2	0*	16*	10*								26	AB.
3	0	16	48	54	82	108					308	O.K.
4	0	18									18	D.A.
5	0	19									19	PR.
6	0	19									19	O.K.
7	0*	20									20	O.K.
8	0	20									20	O.K.
9	0	20	64	189							273	PR.
10	0	21									21	AB.
11	0*	21									21	O.K.
12	0	21									21	PR.
13	0	21	208								229	O.K.
14	0	22	30	28							80	O.K.
15	0	22	22								44	O.K.
16	0	22	26	189	18	25					280	PR.
17	0	23									23	O.K.
18	0*	25									25	O.K.
19	0	25	28	33							86	O.K.
20	0	25	341	44							410	AB.
21	0	25									25	AB.
22	0	25	28	22	69	63	82	271			560	AB.
23	0	26	188								214	O.K.
24	0	27	35	35	29	38	36	42	43	53	338	O.K.
25	0	27									27	AB.
26	0	27									27	PR.
27	0	28									28	O.K.
28	0	28									28	AB.
29	0	29	110	22							161	PR.
30	0	30									30	AB.
31	0	30	31	23	20	20	86	20	54	26	361	AB.
	51										361	AB.
32	0	30	440								470	O.K.
33	0	30									30	AB.
34	0	31	52	31	591						705	O.K.
35	0	31	45								76	AB.
36	0*	31									31	O.K.
37	0	31	105	66							202	O.K.
38	0	31									31	O.K.
39	0	31	98								129	O.K.
40	0	32									32	D.A.
41	0	32									32	AB.
42	0	32	32								64	O.K.
43	0	32									32	O.K.
44	0	33*									33	AB.
45	0	33	35	41	43	57					209	O.K.
46	0	35									35	O.K.
47	0	35	358								393	O.K.
48	0	36	88								124	O.K.
49	0	37	53	40							130	O.K.
50	0	39									39	O.K.
51	0	39									39	O.K.

*Overflow signal.

TABLE II (Cont'd)

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
52	0	40	20	574							634	O.K.
53	0	40*	16	200	432						688	O.K.
54	0	40	409								449	O.K.
55	0	40									40	O.K.
56	0	41	45	55	52	27	25				245	AB.
57	0	41	45	84							170	O.K.
58	0	42	122	68							232	AB.
59	0	43	40	52	38	259					432	AB.
60	0	44									44	O.K.
61	0	46									46	O.K.
62	0	47	32	47	34	40	69				269	O.K.
63	0	47									47	O.K.
64	0	47	179	251							477	O.K.
65	0	48	64								112	AB.
66	0	49									49	AB.
67	0	49	51	57	62	71	60				350	O.K.
68	0	49	96	191							336	O.K.
69	0	50									50	O.K.
70	0	50									50	O.K.
71	0	50	85	151							286	O.K.
72	0	50									50	AB.
73	0	50									50	AB.
74	0	51									51	O.K.
75	0	51									51	AB.
76	0	52									52	O.K.
77	0	52	85	209							346	O.K.
78	0	53	195								248	O.K.
79	0	53									53	AB.
80	0	53									53	O.K.
81	0	55	43								98	AB.
82	0	55	43	27	170*						295	AB.
83	0	56	20	61	36	103					276	AB.
84	0	56	117	57							230	O.K.
85	0	56									56	O.K.
86	0	56	84								140	O.K.
87	0	57	74	81							212	O.K.
88	0	58									58	O.K.
89	0	58	139	84	163	62	127				633	O.K.
90	0	60									60	O.K.
91	0	60	139								199	O.K.
92	0	60									60	O.K.
93	0	60									60	AB.
94	0	61									61	O.K.
95	0	61									61	AB.
96	0	63									63	AB.
97	0	63	31	95	28	20					237	AB.
98	0	64	126	470	85	167					912	O.K.
99	0	64	61	67	84	67					343	O.K.
100	0	64	45	63	63	161					396	O.K.
101	0	65	482								547	O.K.
102	0	66	173	172							411	O.K.
103	0	66†	66	72							204	O.K.
104	0	66									66	AB.

*Overflow signal.

† Don't answer.

TABLE II (Cont'd)

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
105	0	68	66								134	O.K.
106	0	68	330	380							778	O.K.
107	0	69									69	AB.
108	0	70									70	O.K.
109	0	71									71	AB.
110	0	71	95								166	O.K.
111	0	72	112								184	AB.
112	0	72									72	O.K.
113	0	74									74	O.K.
114	0	74	184	93							351	AB.
115	0	75									75	O.K.
116	0	75									75	O.K.
117	0	75	67	203							345	O.K.
118	0	76									76	O.K.
119	0	76									76	AB.
120	0*	77									77	O.K.
121	0	78									78	O.K.
122	0	78									78	O.K.
123	0	78	253	107	38						476	AB.
124	0	79	53								132	O.K.
125	0	80									80	O.K.
126	0	80	50								130	O.K.
127	0	80									80	AB.
128	0	80	117								197	O.K.
129	0	81									81	AB.
130	0	81									81	O.K.
131	0	83									83	O.K.
132	0	84									84	O.K.
133	0	85									85	O.K.
134	0	85	33	294	115						527	O.K.
135	0	88									88	AB.
136	0	88									88	O.K.
137	0*	89									89	O.K.
138	0	90	50	120							260	AB.
139	0	90									90	O.K.
140	0	90									90	AB.
141	0	90	51	39	46						226	AB.
142	0	91	78								169	O.K.
143	0	91									91	O.K.
144	0	91*	48								139	O.K.
145	0	91	116								207	O.K.
146	0*	91									91	O.K.
147	0*	92									92	O.K.
148	0*	92									92	O.K.
149	0	93									93	AB.
150	0	93	34	228	117						472	O.K.
151	0	94	94	75	91						354	AB.
152	0	95									95	AB.
153	0	95									95	O.K.
154	0	97	86	175							358	O.K.
155	0	97	143								240	O.K.
156	0	100									100	O.K.
157	0	100									100	O.K.
158	0	100									100	AB.
159	0	100									100	O.K.
160	0	102	115	198							415	O.K.

*Overflow signal.

TABLE II (Cont'd)

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
161	0	102	80	96	152						430	O.K.
162	0	103									103	O.K.
163	0	104	17								121	O.K.
164	0	105									105	O.K.
165	0	105									105	O.K.
166	0	106	340								446	O.K.
167	0	108	98	140							346	O.K.
168	0	111									111	O.K.
169	0	111									111	AB.
170	0	111	94	125							330	O.K.
171	0	113									113	O.K.
172	0	114									114	O.K.
173	0	116									116	O.K.
174	0	116									116	O.K.
175	0	117									117	O.K.
176	0	120									120	O.K.
177	0	122									122	O.K.
178	0	124	131	209							464	O.K.
179	0	124									124	O.K.
180	0	125	354								479	O.K.
181	0	130									130	O.K.
182	0	130									130	O.K.
183	0	130	125								255	O.K.
184	0	130	56	101							287	O.K.
185	0	131	309								440	O.K.
186	0	134									134	O.K.
187	0	137	147	134	146						564	O.K.
188	0	139	125								264	AB.
189	0	139									139	AB.
190	0	139									139	O.K.
191	0	140	172	60							372	O.K.
192	0	140	400								540	A.B.
193	0	141									141	O.K.
194	0	143									143	O.K.
195	0	143	157								300	O.K.
196	0	144									144	O.K.
197	0	144	187	194	308	115	310	104	165	45	3,463	AB.
	69	90	69	88	87	59	239	277	69	94		
	90	159	193	71	237							
198	0	146									146	O.K.
199	0	146									146	O.K.
200	0	146	184	217							547	AB.
201	0	148									148	O.K.
202	0	149									149	O.K.
203	0	149	28	38	42	46					303	O.K.
204	0	149	121	84							354	A.B.
205	0	150									150	A.B.
206	0	150	26	142	119						437	A.B.
207	0	151	272								423	O.K.
208	0	152	90	95	89	79					505	O.K.
209	0*	155									155	O.K.
210	0	156									156	O.K.
211	0	156									156	O.K.
212	0	156	47	52	217						472	A.B.
213	0	160									160	A.B.
214	0	160									160	O.K.

*Overflow signal.

TABLE II (Cont'd)

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
215	0	160									160	O.K.
216	0	160									160	O.K.
217	0	161									161	A.B.
218	0	164									164	O.K.
219	0	164									164	O.K.
220	0	165									165	A.B.
221	0	168									168	O.K.
222	0	169									169	O.K.
223	0	170									170	O.K.
224	0*	170									170	O.K.
225	0	171									171	O.K.
226	0	175									175	O.K.
227	0	179									179	O.K.
228	0	180									180	O.K.
229	0	181									181	O.K.
230	0	181	360								541	A.B.
231	0	182									182	O.K.
232	0	183									183	O.K.
233	0	183	312	33							528	P.R.
234	0	185									185	O.K.
235	0	186	251								437	A.B.
236	0	192	238								430	A.B.
237	0	195	477								672	O.K.
238	0	198									198	A.B.
239	0	202									202	O.K.
240	0	205	80								285	O.K.
241	0	208									208	O.K.
242	0	209									209	O.K.
243	0	209									209	O.K.
244	0	210									210	O.K.
245	0	214	50	33	29	34	79				439	O.K.
246	0	215	520								735	O.K.
247	0	215									215	A.B.
248	0	217									217	O.K.
249	0*	219									219	D.A.
250	0	219									219	O.K.
251	0	220	163	263	186	123	99	59	105		1,218	AB.
252	0	220	162								382	O.K.
253	0	222									222	O.K.
254	0	226									226	O.K.
255	0	228									228	O.K.
256	0	230									230	AB.
257	0	231	27								258	P.R.
258	0	232									232	O.K.
259	0	235									235	O.K.
260	0*	235									235	O.K.
261	0	238									238	O.K.
262	0	242									242	O.K.
263	0	245									245	O.K.
264	0	246									246	O.K.
265	0	252									252	AB.
266	0	252									252	AB.
267	0	258									258	O.K.
268	0	260	333								593	O.K.
269	0	267	193								460	AB.
270	0	272	219	88*							579	AB.
271	0	278									278	O.K.

*Overflow signal.

TABLE II (Cont'd)

Observation No.	Attempt No.										Total Seconds	Disposition of the Call
	1	2	3	4	5	6	7	8	9	10		
272	0	281	256								281	O.K.
273	0	287									287	O.K.
274	0	288									288	O.K.
275	0	289									545	O.K.
276	0	290									290	O.K.
277	0	296									296	O.K.
278	0	306									306	O.K.
279	0	319									319	O.K.
280	0	320									320	O.K.
281	0	320									320	AB.
282	0*	322	454							322	O.K.	
283	0	331								331	O.K.	
284	0	332								332	DA.	
285	0	338								338	AB.	
286	0	339								339	AB.	
287	0	347								347	O.K.	
288	0	351								805	O.K.	
289	0	351								351	O.K.	
290	0	363								363	O.K.	
291	0	365								365	O.K.	
292	0	369							369	DA.		
293	0	376							376	O.K.		
294	0	378							378	O.K.		
295	0	382							382	O.K.		
296	0	395							395	O.K.		
297	0	398							398	O.K.		
298	0	398							398	AB.		
299	0	400							400	O.K.		
300	0	402							402	O.K.		
301	0	409							409	O.K.		
302	0	416							416	O.K.		
303	0	448							448	O.K.		
304	0	449							449	O.K.		
305	0	455							455	O.K.		
306	0	473							473	O.K.		
307	0	484							484	O.K.		
308	0	484							484	A.B.		
309	0	498							498	O.K.		
310	0	505							505	O.K.		
311	0	509							509	O.K.		
312	0	510							510	A.B.		
313	0	513							513	O.K.		
314	0	526							526	O.K.		
315	0	535	456	541					1,532	O.K.		
316	0	543							543	O.K.		
317	0	556	249						805	O.K.		
318	0	561	389						950	O.K.		
319	0	568							568	O.K.		
320	0	569							569	O.K.		
321	0	570							570	O.K.		
322	0	586							586	O.K.		
323	0	605	(over 600 seconds)						605	A.B.		
324	0	624	(over 600 seconds)						624	A.B.		
325	0	(At 30 seconds received on incoming call from the party desired)										A.B.
326-334	0*	(9 observations)										A.B.
335-451	0	(117 observations)										A.B.

*Overflow signal.

tions that equalled or exceeded particular time intervals between the starts of the first and second attempts. Figure 1(b) shows similar graphical data for 211 Bronx-Westchester observations and Fig. 1(c) shows similar graphical data for 445 Brooklyn-Queens observations. Each of these three graphs is compared with a composite curve for 1107 observations. This composite curve is developed from the data on Fig. 2(a).

Figure 2(a) shows, by dots, the cumulative percentage for 1107 observations, which are comprised of the 451 Manhattan, 211 Bronx-Westchester and 445 Brooklyn-Queens observations, that equalled or exceeded particular time intervals between the starts of the first and second attempts. A smooth curve was drawn through these plotted data. This curve is also shown on other figures, for the purpose of visual comparison of the various plots of data with the overall results.

Figure 2(b) shows a graph concerning 465 observations of the total 1107 observations. These are the cases where a *busy* was observed on a second attempt. (Of the 1107 total observations, 817 resulted in a second attempt within ten minutes and 290 were classified as abandoned. Of the 817 second attempts, 327 cases were able to complete their calls, 16 resulted in a don't answer, 9 were referred to an operator and 465 encountered a *busy*.) Figure 2(b) shows, by dots, the cumulative percentage of the 465 second attempts that equalled or exceeded particular time intervals between the starts of the second and third attempts. The graph of Fig. 2(b) does not differ significantly from the composite curve for 1107 observations. This feature indicates that, when observations concerning subscriber *busies* are made, it is not necessary to have the first observed attempts coincide with the first actual attempts. The observations can begin with any attempt.

Figures 3 and 4 are graphs similar to that shown on Fig. 2(a), the difference being in the graphical ordinates used in order to present additional pictorial representations of the data and to project the curve beyond the observed limits.

The percentage of subscribers who dial their calls again after encountering *busies* is estimated from Figs. 3 and 4 to be 90%. The data on Fig. 3 are projected to a time interval of 1,500 seconds (25 minutes). Judging by eye, beyond this point, it appears that the curve is asymptotic to the 10% horizontal line. This means that 10% of the subscribers abandon their calls and 90% try again. The part of the curve on Fig. 4 that projects beyond the limit of the observed data crosses the 10% line at 6,400 seconds, an interval of $1\frac{3}{4}$ hours. This seems to be a very long time for a subscriber to wait before redialing his call. It is unlikely that many attempts are made beyond this period.

Table III was prepared to determine the disposition of the calls on second attempts and to see if a correlation exists between certain time intervals,

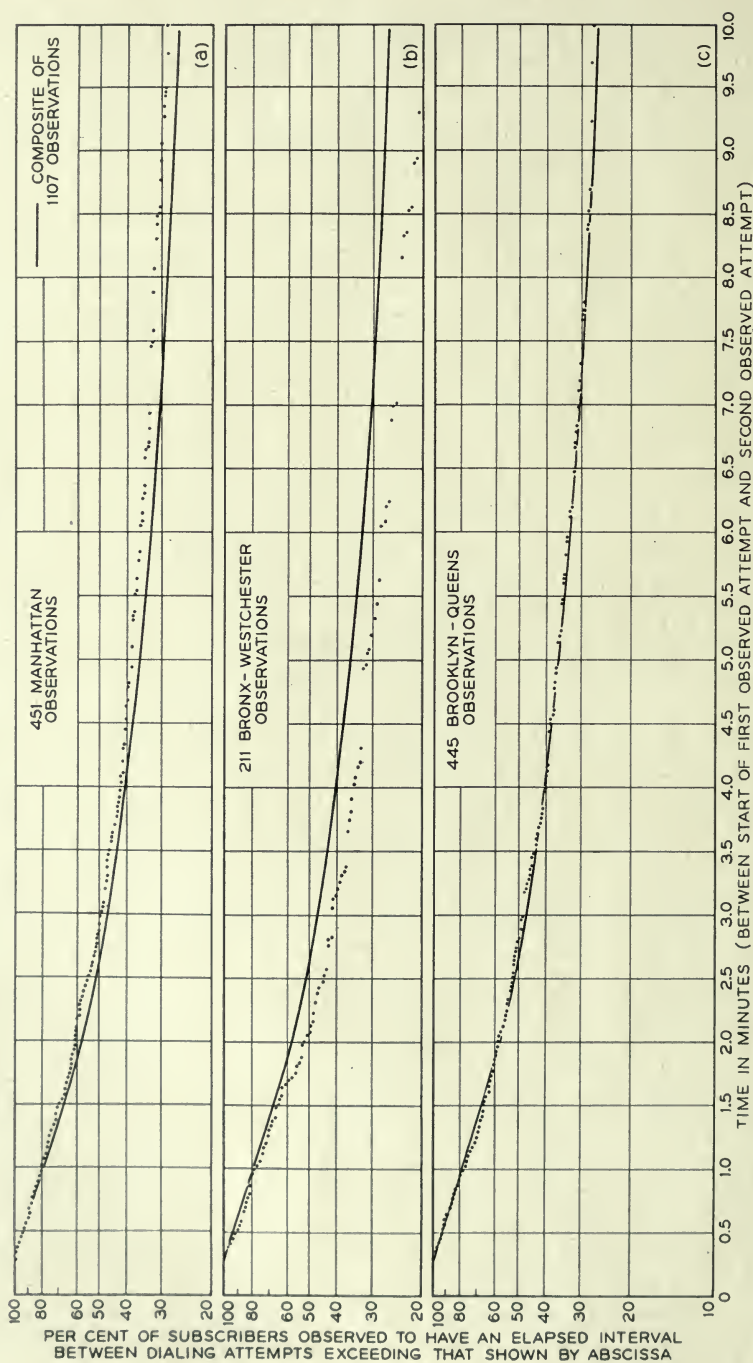


Fig. 1—Results of observations taken at three New York City service observing bureaus concerning the dialing behavior of subscribers upon encountering a busy.

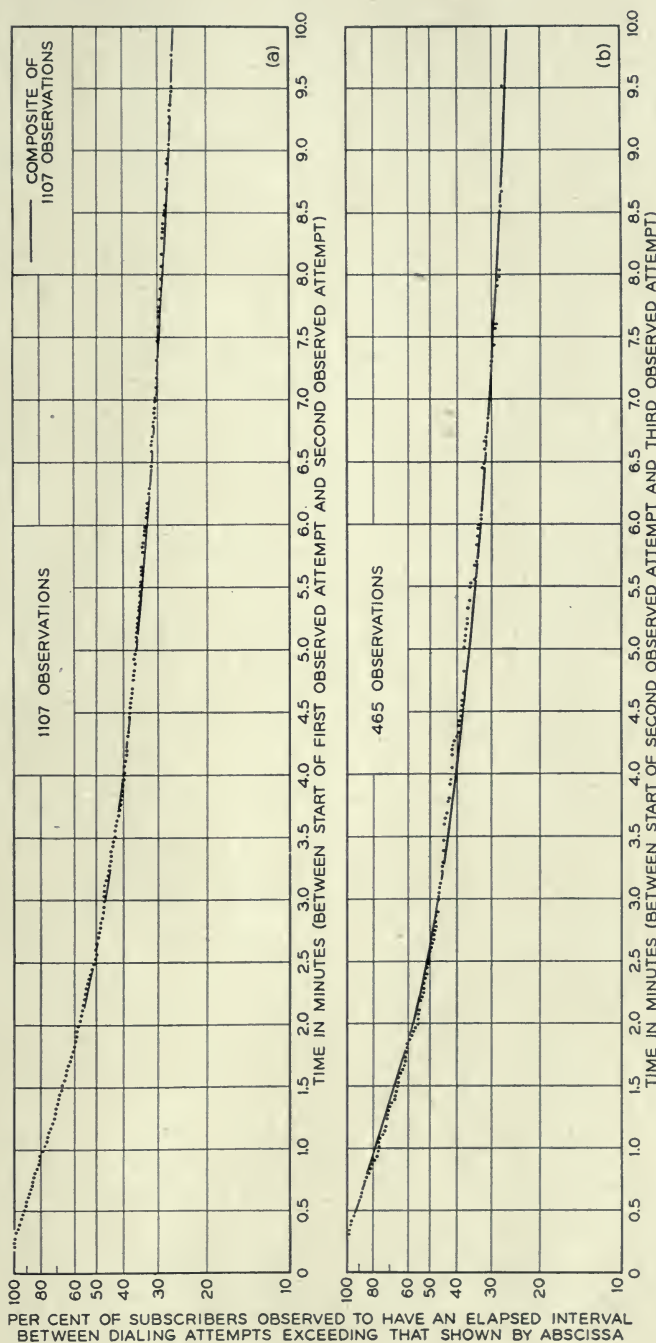
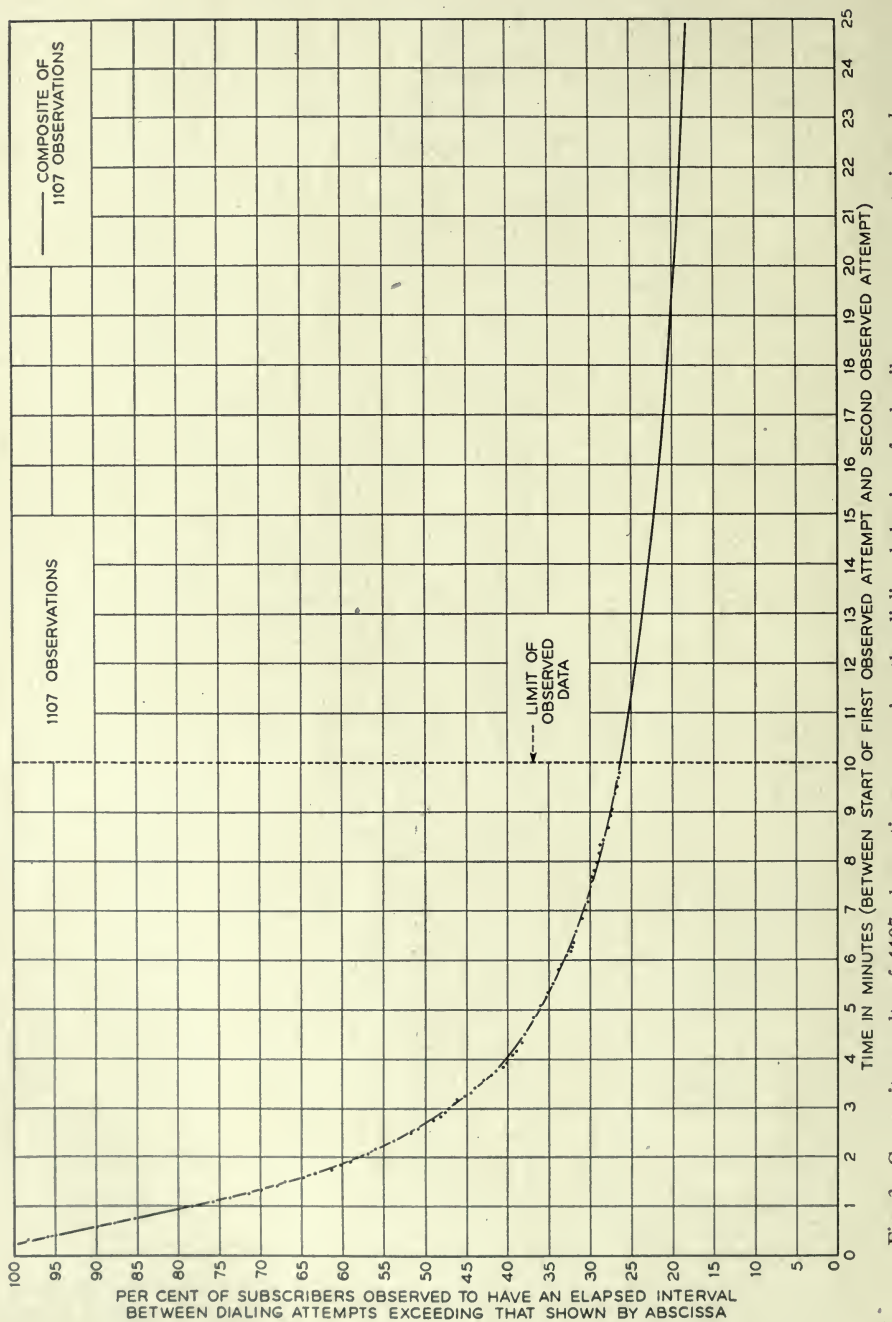


Fig. 2—Composite results of 1107 observations concerning the dialing behavior of subscribers upon encountering a *busy* on a first attempt and results of 465 observations concerning the dialing behavior of subscribers upon encountering a *busy* on a second attempt.



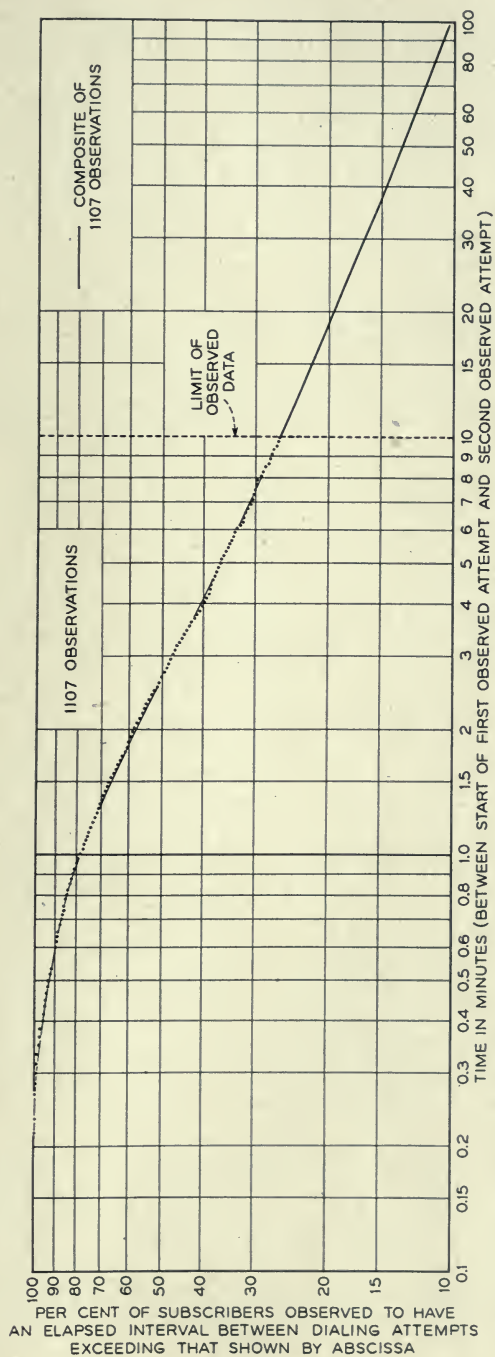


Fig. 4—Composite results of 1107 observations concerning the dialing behavior of subscribers upon encountering a *busy*.

namely, between the first and second attempts and between the second and third attempts. This table was developed by allocating the 817 observations where a second attempt occurred into 5 ranges of time intervals between the first and second attempts of about 163 observations each. For each range of time interval the number of calls that were respectively O.K., DA, PR and AB is listed. Where a third attempt occurred, the numbers of calls are tabulated by ranges of time intervals between the second and third attempts. The ranges of time intervals are the same as

TABLE III
DISPOSITION OF SECOND ATTEMPTS AND CORRELATION OF TIME INTERVALS BETWEEN
DATA CONCERNING 817 OBSERVATIONS HAVING A SECOND ATTEMPT

Range of Time Intervals in Seconds Between the First and Second Attempts	Total Number of Observed Second Attempts	Disposition of the Second Attempt: Number of Second Attempts that				Correlation of Time Intervals Between Attempts: Number of Second Attempts Each of Which Resulted in a <i>Busy</i> and which was Followed by a Third Attempt Within the Range of Seconds Listed in the Column Headings Below				
		Were OK	Were DA	Were PR	Were AB	0-45	46-78	79-130	131-226	227-600
0- 45	164	36	7	5	26	43*	17	12	8	10
46- 78	164	44	1	2	24	14	26*	28	13	12
79-130	164	71	2	2	21	6	14	24*	11	13
131-226	162	83	2	0	27	5	9	5	15*	16
227-600	163	93	4	0	28	6	0	4	5	23*
	817	327	16	9	126	74	66	73	52	74

* The asterisk marks the items that had the same range of time intervals between the first and second attempts and between the second and third attempts.

those used between the first and second attempts in order to see if a correlation exists. The significant facts concerning these data are:—

1. The degree of success in obtaining an O.K. call was better for those subscribers who waited longer before making a subsequent attempt. Only 22% of the subscribers who waited from 0 to 45 seconds were successful as against 57% who waited from 227 to 600 seconds.
2. The number of calls referred to the operator or where don't answers occurred are not significant to the problem in hand.
3. The incidence of abandoned calls appears to be uniform for the five ranges of time intervals. This means that the 90% figure estimated from Fig. 3 can be considered to apply with equal effect to all subscribers without regard to the previous time interval between dialing attempts.
4. The correlation data indicate a tendency for subscribers to establish a tempo or pace which they follow when redialing their calls. If this tempo did not exist the items on Table III that are marked with asterisks would not be larger than the surrounding items.

It was previously indicated that no observations were taken on P.B.X. and coin lines. An earlier attempt to collect data concerning the behavior of subscribers when encountering *busies* produced data that showed fewer subsequent attempts than was believed to be the case. The differences between the earlier data, which included a high proportion of observations on P.B.X. and coin lines, and the data developed herein are believed to be fully

accounted for and it is believed that the P.B.X. and coin lines have the same basic characteristics regarding dialing behavior upon encountering *busies* as have the subscribers who were observed. No significant differences between the results for residential and business offices were noted. From these indirect facts, it is concluded that no significant differences exist between classes of subscribers.

EFFECT ON THE TRUNK PLANT

As explained earlier, neither the Poisson nor the Erlang B formula gives an accurate picture of the facts when trunk shortages occur on trunk groups handling subscriber-dialed calls. In both formulae it is assumed that only one attempt is made per call. In the case of the Poisson formula, the call is assumed to be held by the dial equipment until a trunk becomes available or until the subscriber hangs up, and in the case of the Erlang B formula, the call is assumed to clear out. The data developed from the service observations, concerning the dialing behavior of subscribers when encountering *busies*, indicate that subscribers usually make many subsequent attempts when a *busy* is encountered. Also the dial equipment with which we are familiar clears out the calls by giving an *all-trunks-busy* signal. In order to determine what a trunk capacity table might be like that takes into account the habits of subscribers and the limitations of the dial equipment a study based on simulated traffic was made. This study consisted of 150 CCS (hundred call seconds per hour) of traffic offered to a trunk group varying from 5 to 12 trunks. This study utilized the data developed from the service observations.

A study based on simulated traffic is a method used to study the capacities of trunking arrangements where a formula is not available. This type of study is based on the idea that calls are placed at random, that holding times of the calls follow an exponential law, and that these characteristics can be simulated by random numbers drawn from an appropriate source.

The study of 150 CCS of simulated traffic was based on 1,000 calls offered to a trunk group during a ten-hour period. The average holding time per call was 150 seconds, with the total holding time being 150,000 seconds or 41.66667 hours. Sub-divisions of an hour were expressed in decimal terms, the smallest division being a hundred-thousandth part. Three sets of random numbers were used for the following purposes:

1. To determine at what time in the ten-hour period a particular call is offered to the trunk group.
2. To furnish the holding time of a particular call.
3. To define for each call the pattern of resubmission of the call to the trunk group should an *all-trunks-busy* be encountered by the call.

In each instance the numbers were taken from the tail-end portions of

successive entries of 19 significant figures of e^x (Tables of the Exponential Function—WPA—1939). The numbers drawn and their functions in the study are as follows:

A set of 1,000 six-digit numbers was taken from the last six digits of entries of e^x from $x = 0.4000$ to $x = 0.4999$. These 1,000 six-digit numbers were arranged in numerical order to give the placing time of 1,000 simulated calls. The first digit in every number was used to represent the hour and the last five digits the hundred-thousands part of the hour when a particular call was placed. The randomness of this particular draw was checked by determining the differences between successive placement times and then arranging the differences in numerical order. The results were plotted on a cumulative basis on Fig. 5, where a visual comparison can be made with theoretical results.

A set of 1,000 seven-digit random numbers between 0,000,000 to 4,166,667 inclusive were taken from the last seven digits of entries of e^x from $x = 0.5000$ to $x = 0.7344$. Numbers above 4,166,667 were disregarded. These seven-digit numbers when arranged in numerical order accounted for the total holding time of all the calls. The difference between successive numbers arranged in numerical order, furnished 1,000 individual holding times.

A third set of 1,000 random numbers were taken from two sources in the e^x tables. These 1,000 numbers contained a variable number of digits. These numbers were for use when calls encountered *all trunks busies* in order to determine which calls were to be resubmitted and to determine the time interval for resubmitting a call. Previously, it was estimated from Fig. 3, that 90% of the subscribers after encountering a *busy* redial their call. This estimate was used by assigning to the numerals 1 to 9 in the third set of random numbers the characteristic that a call may make a subsequent attempt if it encounters an *all trunks busy* and by assigning to the numeral 0 the characteristic that the call drops out if it encounters an *all trunks busy*. About 10% of the 1,000 numbers show a numeral 0 in the first place and hence no further digits are needed because the call drops out. The remaining 90% of the numbers show numerals from 1 to 9 in the first place and hence may make a second attempt. If an *all trunks busy* is encountered on the second attempt, a numeral from 1 to 9 in the second place determines that a third attempt may be made while the numeral 0 determines that the call drops out. This process is repeated for each place of each number in the third set of 1,000 random numbers until the numeral 0 appears. The number of consecutive places showing only numerals from 1 to 9, indicates the total number of attempts that a particular call might make before it drops out. Thus for a particular number the numerals might be 4720. In this case, three subsequent attempts can be made. Another number might be 834650. In this case, five subsequent attempts can be made.

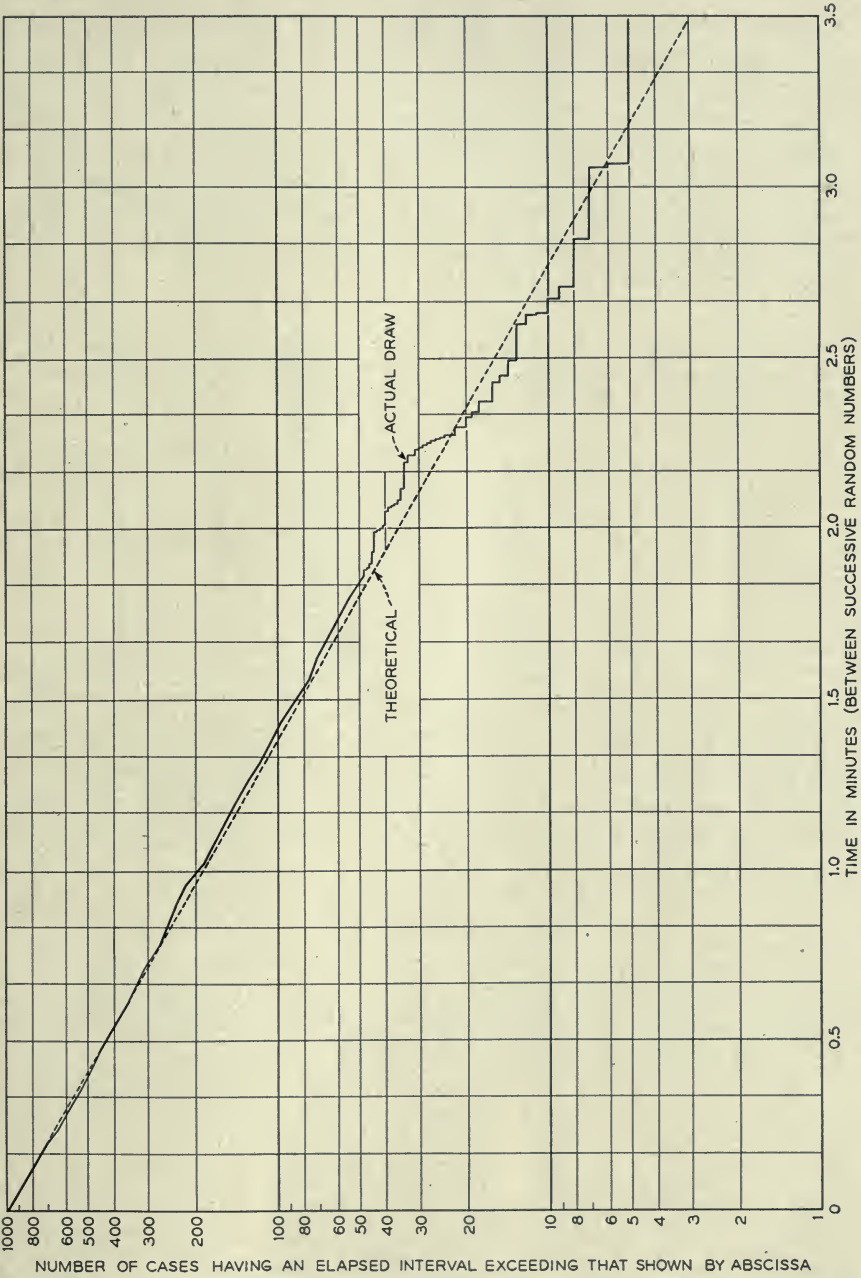


Fig. 5—Check of the randomness of numbers drawn that represent the time of placement of calls.

The effect of using numerals in this way is that 90% of the calls encountering *all trunks busies* appear as subsequent attempts.

The numeral in the first place of each of the third set of random numbers was used to establish the time interval for resubmitting each call. The time intervals were developed from the data on Fig. 4 by dividing the vertical scale into 10% bands. The time interval corresponding to the midpoint of each band was used as applicable to the 10% of the calls that fell within that band. The midpoint values, the corresponding time intervals, and the random numerals used are as follows:

TABLE NO. IV

Midpoint Values of the 10% Bands of figure 4	Corresponding Time Intervals in Seconds	Equivalent Hundred-Thousandth Part of an Hour	Assignment of Random Numerals
<i>a</i>	<i>b</i>	$c = b \div .036$	<i>d</i>
95	25	700	9
85	46	1,300	8
75	67	1,900	7
65	93	2,600	6
55	132	3,700	5
45	195	5,400	4
35	320	8,900	3
25	665	18,500	2
15	2,250	62,500	1
5	Infinite	Call drops out	0

Based on the results indicated by Table III, that subscribers tend to make repetitious attempts at a uniform pace or tempo, the time interval determined by the numeral in the first place of a particular number of the third set of random numbers was repeated each time that a particular call was resubmitted.

The results of the study of simulated traffic are as follows:

TABLE V

Trunks Provided	Attempts (Calls Offered Plus All Subsequent Attempts)	Overflows (Calls Encountering All Trunks Busies)	Ratios of Overflows to Attempts	Calls Handled	Calls Abandoned	Approx. CCS Handled
<i>a</i>	<i>b</i>	<i>c</i>	$d = c \div b$	$e = b - c$	$f = 1000 - e$	$g = .150xe$
5	1,658	720	.4343	938	62	141
6	1,287	319	.2479	968	32	145
7	1,147	155	.1351	992	8	149
8	1,071	75	.0700	996	4	149
9	1,027	28	.0273	999	1	150
10	1,011	12	.0119	999	1	150
11	1,005	5	.0050	1,000	0	150
12	1,000	0	.0000	1,000	0	150

The ratios of overflows to attempts compared with theoretical results for the Poisson and Erlang B formulae for 150 CCS of offered traffic are as follows:

TABLE VI

Trunks Provided	Study of Simulated Traffic: Ratios of Overflows to Attempts	Theoretical Results	
		Erlang B: Ratio of Calls Lost to Calls Offered	Poisson: Ratio of Calls Delayed to Calls Offered
5	.4343	.2139	.4037
6	.2479	.1293	.2414
7	.1351	.0715	.1288
8	.0700	.0359	.0617
9	.0273	.0163	.0268
10	.0119	.0068	.0106
11	.0050	.0026	.0038
12	.0000	.0009	.0013

The ratios of overflows to attempts are apparently very close to the Poisson results. No further conclusion should be drawn from this, at this time, without further study.

SUMMARY

Data concerning the dialing behavior of subscribers who encounter *busies* have been obtained for New York City subscribers. These data indicate quantitatively: (1) how soon after obtaining a *busy*, a subscriber redials his call; (2) what percentage of subscribers make subsequent attempts; and (3) the pattern of time intervals between successive subsequent attempts. These data appear to have direct application in the development of trunk capacity tables for trunks handling subscriber-dialed traffic when trunk shortages occur.

ACKNOWLEDGMENTS

The writer gratefully acknowledges the help of Mr. H. P. Penny in planning the method for taking the service observations and in taking the Manhattan and Bronx-Westchester observations. The help of Mr. R. A. Colbeth is acknowledged in taking the Brooklyn-Queens observations.

Spectra of Quantized Signals

By W. R. BENNETT

1. DISCUSSION OF PROBLEM AND RESULTS PRESENTED

SIGNALS which are quantized both in time of occurrence and in magnitude are in fact quite old in the communications art. Printing telegraph is an outstanding example. Here, time is divided into equal divisions, and the number of magnitudes to be distinguished in any one interval is usually no more than two, corresponding to the closed or open positions of a sending switch. It is only in recent years, however, that the development of high speed electronic devices has progressed sufficiently to enable quantizing techniques to be applied to rapidly changing signals such as produced by speech, music, or television. Quantizing of time, or time division, has found application as a means of multiplexing telephone channels.¹ The method consists of connecting the different channels to the line in sequence by fast moving switches synchronized at the transmitting and receiving ends. In this way a transmission medium capable of handling a much wider band of frequencies than required for one telephone channel can be used simultaneously by a group of channels without mutual interference. The plan is the same as that used in multiplex telegraphy. The difference is that ordinary rotating machinery suffices at the relatively low speeds employed by the latter, while the high speeds needed for time division multiplex telephony can be realized only by practically inertialess electron streams. Also the widths of frequency band required for multiplex telephony are enormously greater than needed for the telegraph, and in fact have become technically feasible only with the development of wide-band radio and cable transmission systems. As far as any one channel is concerned the result is the same as in telegraphy, namely that signals are received at discrete or quantized times. In the limiting case when many channels are sent the speech voltage from one channel is practically constant during the brief switch closure and, in effect, we can send only one magnitude for each contact or quantum of time. The more familiar word "sampling" will be used here interchangeably with the rather formidable term "quantizing of time".

Quantizing the magnitude of speech signals is a fairly recent innovation. Here we do not permit a selection from a continuous range of magnitudes but only certain discrete ones. This means that the original speech signal

is to be replaced by a wave constructed of quantized values selected on a minimum error basis from the discrete set available. Clearly if we assign the quantum values with sufficiently close spacing we may make the quantized wave indistinguishable by the ear from the original. The purpose of quantization of magnitudes is to suppress the effects of interference in the transmission medium. By the use of precise receiving instruments we can restore the received quanta without any effect from superposed interference provided the interference does not exceed half the difference between adjacent steps.

By combining quantization of magnitude and time, we make it possible to code the speech signals, since transmission now consists of sending one of a discrete set of magnitudes for each distinct time interval.^{2,3,4,5,6,7} The maximum advantage over interference is obtained by expressing each discrete signal magnitude in binary notation in which the only symbols used are 0 and 1. The number which is written as 4 in decimal notation is then represented by 100, 8 by 1000, 16 by 10,000; etc. In general, if we have N digit positions in the binary system, we can construct 2^N different numbers. If we need no more than 2^N different discrete magnitudes for speech transmission, complete information can be sent by a sequence of N on-or-off pulses during each sampling interval. Actually a total of $2^N!$ different coding plans (sets of one-to-one correspondences between signal magnitudes and on-or-off sequences) is possible. The straightforward binary number system is taken as a representative example convenient for either theoretical discussion or practical instrumentation. We assume that absence of a pulse represents the symbol 0 and presence of a pulse represents the symbol 1. The receiver then need only distinguish between two conditions: no transmitted signal and full strength transmitted signal. By spacing the repeaters at intervals such that interference does not reach half the full strength signal at the receiver, we can transmit the signal an indefinitely great distance without any increment in distortion over that originally introduced by the quantizing itself. The latter can be made negligible by using a sufficient number of steps.

To determine the number of quantized steps required to transmit specific signals, we require a knowledge of the relation between distortion and step size. This problem is the subject of the present paper.* We divide the problem into two parts: (1) quantizing the magnitude only and (2) combined quantizing of magnitude and time. The first part can be treated by a simple model: the "staircase transducer", which is a device having the instantaneous output vs. input curve shown by Fig. 1. Signals impressed on the stair-

* Other features of the quantizing and coding theory are discussed in forthcoming papers by Messrs. C. E. Shannon, J. R. Pierce, and B. M. Oliver.

case transducer are sorted into voltage slices (the treads of the staircase), and all signals within plus or minus half a step of the midvalue of a slice are replaced in the output by the midvalue. The corresponding output when the input is a smoothly varying function of time is illustrated in Fig. 2. The output remains constant while the input signal remains within the boundaries of a tread and changes abruptly by one full step when the signal crosses the boundary. It is not within the scope of the present paper to discuss the internal mechanism of a staircase transducer, which may have many different physical embodiments. We are concerned rather with the distortion produced by such a device when operating perfectly.

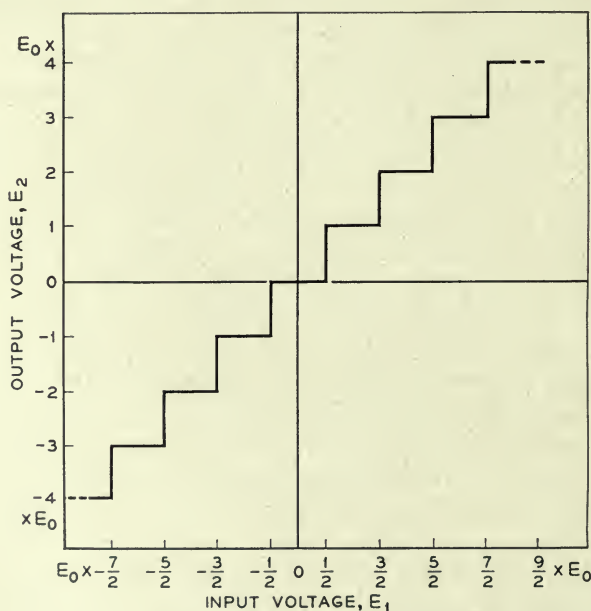


Fig. 1—Quantizing characteristic.

The distortion or error consists of the difference between the input and output signals. The maximum instantaneous value of distortion is half of one step, and the total range of variation is from minus half a step to plus half a step. The error as a function of input signal voltage is plotted in Fig. 3 and a typical variation with time is indicated in Fig. 2. If there is a large number of small steps, the error signal resembles a series of straight lines with varying slopes, but nearly always extending over the vertical interval between minus and plus half a step. The exceptional cases occur when the signal goes through a maximum or minimum within a step. The limiting condition of closely spaced steps enables us to derive quite simply

an approximate value for the mean square error, which will later be shown to be sufficiently accurate in most cases of practical importance. This approximation consists of calculating the mean square value of a straight line going from minus half a step to plus half a step with arbitrary slope. If

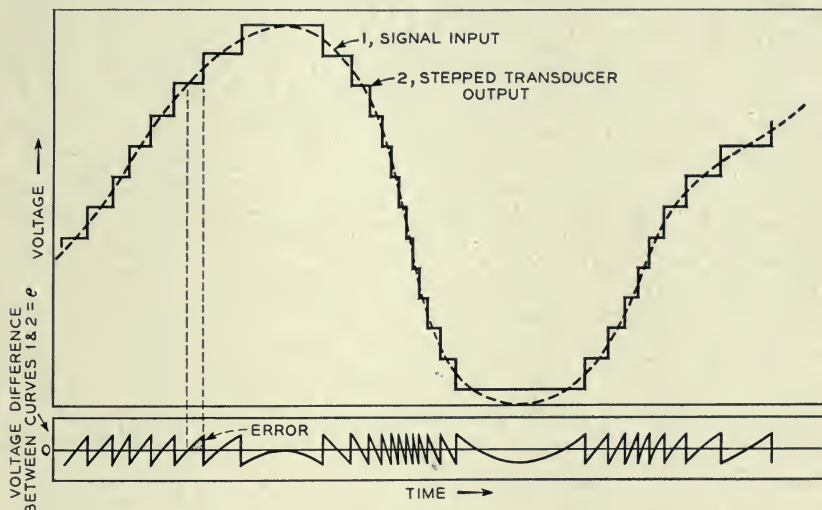


Fig. 2—A quantized signal wave and the corresponding error wave.

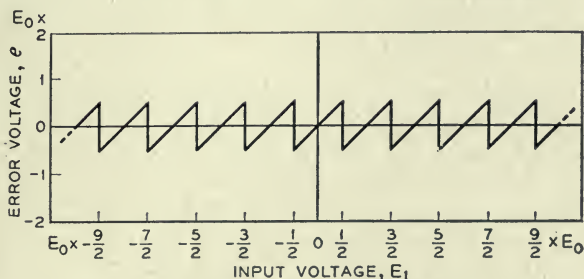


Fig. 3—Characteristic of the errors in quantizing.

E_0 is the voltage corresponding to one step, and s is the slope, the equation of the typical line is:

$$\epsilon = st, \quad -\frac{E_0}{2s} < t < \frac{E_0}{2s} \quad (1.0)$$

where ϵ is the error voltage and t is the time referred to the midpoint as origin. Then the mean square error is

$$\bar{\epsilon}^2 = \frac{s}{E_0} \int_{-E_0/2s}^{E_0/2s} \epsilon^2 dt = \frac{E_0^2}{12}, \quad (1.1)$$

or one twelfth the square of the step size.

Not all the distortion falls within the signal band. The distortion may be considered to result from a modulation process consisting of the application of the component frequencies of the original signal to the non-linear staircase characteristic. High order modulation products may have frequencies quite remote from those in the original signal and these can be excluded by a filter passing only the signal band. It becomes of importance, therefore, to calculate the spectrum of the error wave. This we shall do in the next section for a generalized signal using the method of correlation, which is based on the fact that the power spectrum of a wave is the Fourier cosine transform of the correlation function. The result is then applied to a particular kind of signal, namely one having energy uniformly distributed throughout a definite frequency band and with the phases of the components randomly distributed. This is a particularly convenient type of signal because it in effect averages over a large number of possible discrete frequency components within the band. Single or double-frequency signal waves are awkward for analytical purposes because of the ragged nature of the spectra produced. The amplitudes of particular harmonics or cross-products of discrete frequency components are found to oscillate violently with magnitude of input. The use of a large number of input components smooths out the irregularities.

The type of spectra obtained is shown in Fig. 4. Anticipating binary coding, we have shown results in terms of the number of binary digits used. The number of different magnitudes available are 16, 32, 64, 128, and 256 for $N = 4, 5, 6, 7$ and 8 digits, respectively. Here a word of explanation is needed with respect to the placing of the scale of quantized voltages. A signal with a continuous distribution of components along the frequency scale is theoretically capable of assuming indefinitely great values of instantaneous voltage at infrequent instants of time. An actual quantizer (staircase transducer) has a finite overload value which must not be exceeded and hence can have only a finite number of steps. This difficulty is resolved here by the experimentally observed fact that thermal noise, which has the type of spectrum we have assumed for our signal, has never been observed to exceed appreciably a voltage four times its root-mean-square value. Hence we have placed the root-mean-square value of the input signal at one-fourth the overload input to the staircase. This fixes the relation between step size and the total number of steps. In the actual calculation the number of steps is taken as infinite; the effect of the assumed additional steps beyond 2^N is negligible because of the rarity of excursion into this range.

The curves of Fig. 4 are drawn for the case in which the signal band starts at zero frequency. The original signal band width is represented by one unit on the horizontal scale. The relatively wide spread of the distortion spectrum is clearly shown. As the number of digits (or steps) is increased

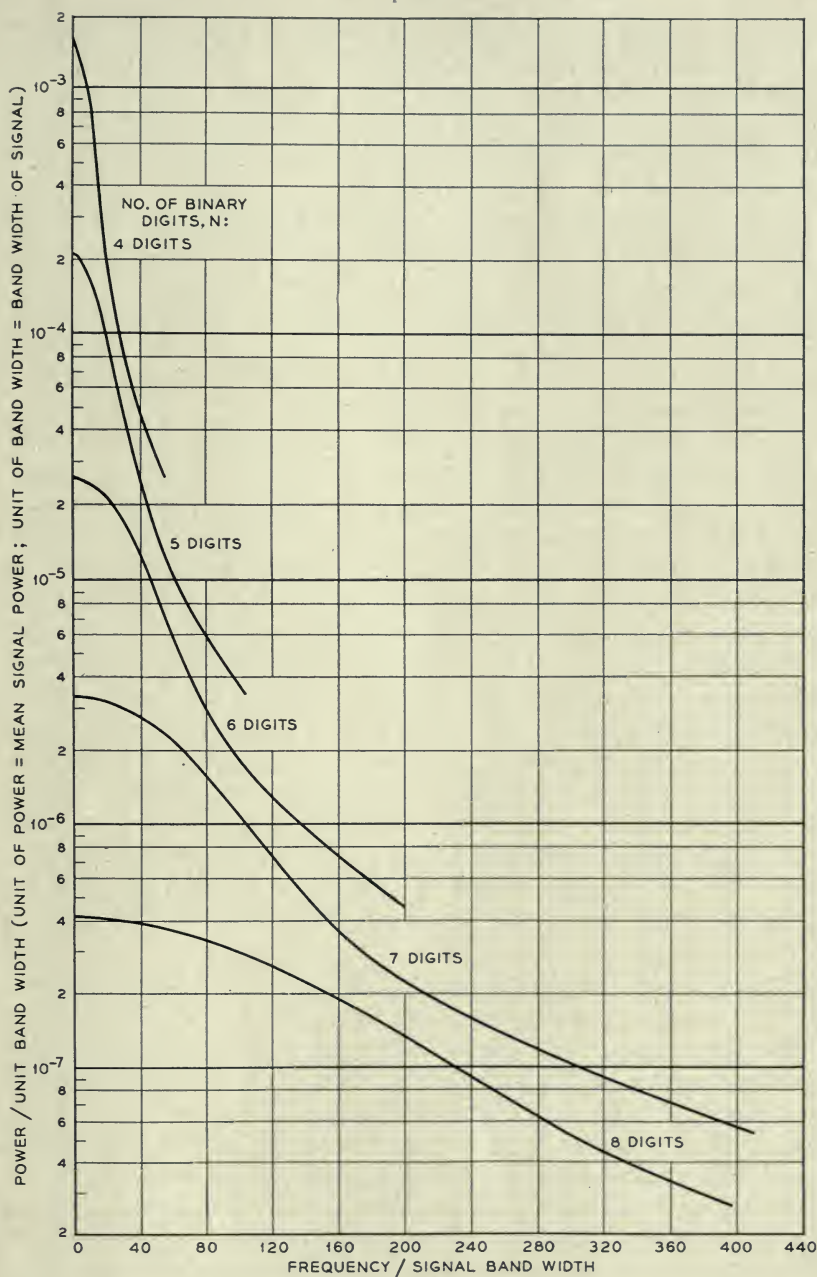


Fig. 4—Spectrum of distortion from quantizing the magnitudes of a random noise wave. Full load on the quantizer is reached by peaks 12 db above the r.m.s. value of input.

the spectrum becomes flatter over a wider range, but with a smaller maximum density. The area under each curve represents the total mean power in the corresponding error wave and is found to agree quite accurately with the approximate result of Eq. (1.1). The distortion power falling in the signal band is represented by the area included under the curve from zero to unit abscissa.

Quantizing the magnitude only is not a technically attractive method of transmission because of the wide frequency band required to preserve the discrete values of the quanta. Thus in a 128-step system, a full load sinusoidal signal passes through 64 different steps each quarter cycle and hence would require transmitting 256 successively different magnitudes during each period of the signal frequency. We therefore consider the second problem—that of sampling the quantized magnitudes.

The theory of periodic sampling of signals is a limiting case of commutator modulation theory as previously shown by the author.¹ We may think of a periodically closed switch in series with the line and source as producing a multiplication of the signal by a switching function. The switching function has a finite value during the time of switch closure and is zero at other times. It may be expanded in a Fourier series containing a term of zero frequency, the repetition frequency of switch closure, and all harmonics of the latter. Multiplication of the signal by the Fourier series representing the constant component of the switching function gives a term proportional to the signal itself. Multiplication of the signal by the fundamental component of the switching function gives upper and lower sidebands on the repetition frequency. Likewise multiplication by the harmonics gives sidebands on each harmonic. The signal is separable from the sidebands on a frequency basis if the signal band does not overlap the lower sideband on the repetition frequency. This leads to the condition for no distortion in time division: the highest signal frequency must be less than one-half the repetition frequency.

To apply the above theory to instantaneous sampling we let the duration of switch closure in one period approach zero. We then approach the condition of one signal value in each period, so that the repetition frequency now becomes the sampling frequency. Clearly the sampling frequency must slightly exceed twice the highest signal frequency. We also note that as the contact time tends toward zero, the switching function approaches a periodically repeated impulse. The important terms of the Fourier series representing the switching function accordingly become a set of harmonics of equal amplitude with a constant component equal to half the amplitude of the typical harmonic. On multiplication of this series by the signal, we get a set of sidebands of equal amplitude including the one corresponding to the original signal itself, the sideband on zero frequency.

These results may be applied to the staircase transducer. The output may be resolved into the input signal plus the error. The sampling frequency is assumed to exceed its minimum required value of twice the top signal frequency. The component of the output that is equal to the original signal can therefore be separated at the receiver by a filter passing the original signal band. A similar statement cannot be made for the error component, for it has been found to extend over a vastly greater range than the original signal. To calculate the total distortion received in the signal band, we can multiply the distortion spectrum by the switching function and sum up all sideband contributions to the original signal band. Each har-

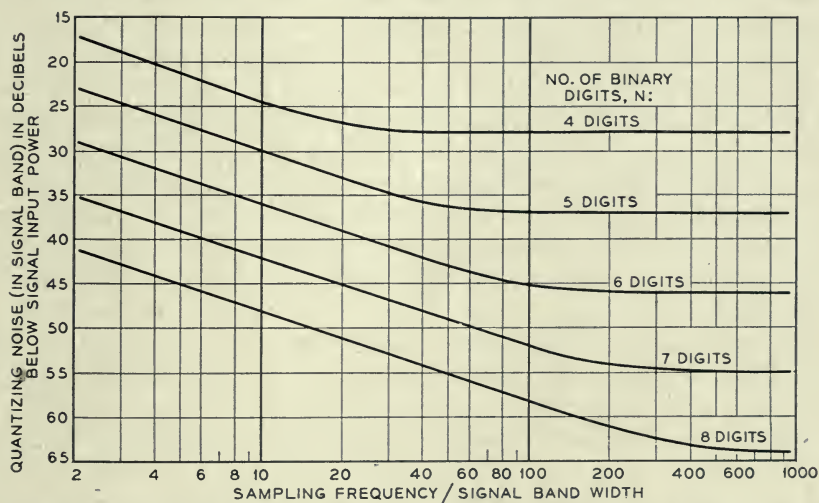


Fig. 5—Total distortion in signal band from quantizing and sampling a random noise wave. Full load on the quantizer is 12 db above the r.m.s. value of input.

monic of the switching function makes such contributions by beating with a band of the error spectrum above and below the frequency of the harmonic. These contributions add as power when the sampling frequency is independent of the individual frequencies contained in the signal. The total error power accepted by the signal band filter decreases as the sampling frequency is increased because each harmonic of the sampling frequency is thereby pushed upward into a less dense portion of the error spectrum. In the limit as the sampling frequency is made indefinitely large, we return to the non-sampled case, that of the staircase transducer only.

Figure 5 shows the calculated curves of distortion in the signal band plotted as a function of ratio of sampling frequency to signal band width. The curves have downward slopes approaching asymptotes corresponding to the area from zero to unity under the corresponding curves of Fig. 4.

The initial points at the minimum sampling rate are determined on the other hand by the total area under the curves of Fig. 4, since the accepted sidebands on the harmonics in this case exactly fill out the entire error spectrum. These initial points are therefore given quite accurately by Eq. (1.1), which, as pointed out before, is a good approximation for the total areas. We can also give a direct demonstration of the applicability of Eq. (1.1) to the initial points of the curves of Fig. 5 by means of the following theorem:

Theorem I. The mean square value of the response of an ideal low-pass filter to a train of unit impulses multiplied by instantaneous samples occurring at double the cutoff frequency is equal to the mean square value of the samples provided no harmonic of the sampling frequency is equal to twice the frequency of one component or equal to the sum or difference of two component frequencies of the sampled signal. Proof of the theorem is given in Appendix I. To apply it here we resolve the input into two components: the true signal and the error. The former is reproduced with fidelity in the output because it contains only frequencies below half the sampling rate. The error component in the output represents the response of the low-pass filter to the error samples. Except for very special types of signals, the error samples are uniformly distributed throughout the range from minus half a step to plus half a step. Calculation of the mean square value of such a distribution gives Eq. (1.1).

We have tacitly assumed above that the sampled values applied to the filter in the output of the system are infinitesimally narrow pulses of height proportional to the samples. In actual systems it is found advantageous to hold the sampled values constant in the individual receiving channels until the next sample is received. This means that the input to the channel filter is a succession of rectangular pulses of heights proportional to the samples. The resulting magnitude of recovered signal is much larger than would be obtained if very short pulses of the same heights were used; stretching the pulses in time produces in effect an amplification. The amplification is obtained, however, at the expense of a variation of channel transmission with signal frequency. Infinitesimally short pulses have a flat frequency spectrum, while pulses of finite duration do not. The frequency characteristic introduced by lengthening the pulses is easily calculated by determining the steady state admittance function of a network which converts impulses to the actual pulses used. The general formula for this admittance when a unit impulse input is converted into an output pulse $g(t)$ is easily shown to be:

$$Y(i\omega) = f_s \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt \quad (1.2)$$

where f_s is the repetition frequency and ω is the angular signal frequency.

We shall call this Theorem II and give the proof in Appendix II. This relation is similar to that found in television and telephotography for the "aperture effect", or variation of transmission with frequency caused by the finite size of the scanning aperture. The pulse shape $g(t)$ is analogous to a variation in aperture height $g(x)$, where x is distance along the line of scanning. Hence it has become customary to use the term "aperture effect" in the theory of restoring signals from samples. The aperture effect associated with rectangular pulses lasting from one sample to the next amounts to an amplitude reduction of $\pi/2$ or 3.9 db at the top signal frequency (one half the sampling rate) compared to a signal of zero frequency. There is also a constant delay introduced equal to half the sampling period. The latter does not cause any distortion and the amplitude effect can be corrected by properly designed equalizing networks.

The fact that many pulse spectra can be simply expressed in terms of a flat spectrum associated with sharp pulses and an aperture effect caused by the particular shape of pulse used does not appear to have been recognized in the recent literature, although applications were made by Nyquist in a fundamental paper⁸ of 1928. Premature introduction of a specific finite pulse not only complicates the work, but also restricts the generality of the results.

Distortion caused by quantizing errors produces much the same sort of effects as an independent source of noise. The reason for this is that the spectrum of the distortion in the receiving filter output is practically independent of that of the signal over a wide range of signal magnitudes. Even when the signal is weak so that only a few quantizing steps are operated, there is usually enough residual noise on actual systems to determine the quantizing noise and mask the relation between it and the signal. Eq. (1.1) yields a simple rule enabling one to estimate the magnitude of the quantizing noise with respect to a full load sine wave test tone. Let the full load test tone have peak voltage E ; its mean square value is then $E^2/2$. The total range of the quantizer must be $2E$ because the test signal swings between $-E$ and $+E$. The ratio $2E/E_0 = r$ is a convenient one to use in specifying the quantizing; it is the ratio of the total voltage range to the range occupied by one step. The ratio of mean square signal to mean square quantizing noise voltage is

$$\frac{E^2/2}{E_0^2/12} = \frac{6E^2}{4E^2/r^2} = \frac{3r^2}{2} \quad (1.3)$$

Actual systems fail to reproduce the full band $f_s/2$ because of the finite frequency range needed for transition from pass-band to cutoff. If we introduce a factor κ to represent the ratio of equivalent rectangular noise band

to $f_s/2$, the actual received noise power is multiplied by κ . Then the signal-to-noise ratio in db for a full load test tone is

$$D = 10 \log_{10} \frac{3r^2}{2\kappa} \text{ db} \quad (1.4)$$

In practical applications the value of κ is about $3/4$ which gives the convenient rule:

$$D = 20 \log_{10} r + 3 \text{ db} \quad (1.5)$$

In other words, we add 3 db to the ratio expressed in db of peak-to-peak quantizing range to the range occupied by one step. For various numbers of binary digits the values of D are:

TABLE I

Number of Digits	D
3	21
4	27
5	33
6	39
7	45
8	51

From Table I we can make a quick estimate of the number of digits required for a particular signal transmission system provided that we have some idea of the required signal-to-noise ratio for a full load test tone. The latter ratio may be expressed in terms of the full load test tone which the system is required to handle and the maximum permissible unweighted noise power at the same level point. Since quantizing noise is uniformly distributed throughout the signal band, its interfering effect on speech or other program material is probably similar to that of thermal noise with the same mean power. Requirements given in terms of noise meter readings must be corrected by the proper weighting factor before applying the table. If the signal transmitted is itself a multiplex signal with channels allotted on a frequency division basis, the noise power falling in each channel is the same fraction of the total noise power as the band width occupied by the signal is of the total band width of the system.

We have thus far considered only the case in which the quantized steps are equal. In actual systems designed for transmission of speech it is found advantageous to taper the steps in such a way that finer divisions are available for weak signals. For a given number of total steps this means that coarser quantization applies near the peaks of large signals, but the larger absolute errors are tolerable here because they are small relative to the bigger signal values. Tapered quantizing is equivalent to inserting complementary non-linear transducers in the signal branch before and after the quantizer. In

the usual case, the transducer ahead of the quantizer is of the "compressing" type in which the loss increases as the signal increases. If the full load signal just covers all the linear quantizing steps, a weak signal gets a bigger share of the steps than it would if the transducer were linear. The transducer after the quantizer must be of the "expanding" type which gives decreased loss to the large signals to make the overall combination linear.

On the basis of the theory so far discussed, we can say that the error spectrum out of the linear quantizer is virtually the same whether or not the signal input is compressed. The operation of the expander then magnifies the errors produced when the signal is large. When weak signals are applied, the mean square error is given by Eq. (1.1), as before, but when the signal is increased an increment in noise occurs. The mean square value of noise voltage under load may be computed from the probability density of the signal values and the output-vs-input characteristic of the expander, or its inverse, the compressor. A first order approximation, valid when the steps are not too far apart, replaces (1.1) by:

$$\overline{\epsilon^2} = \frac{E_0^2}{12} \int_{Q_2}^{Q_1} \frac{p_1(E_1) dE_1}{[F'(E_1)]^2} \quad (1.6)$$

where Q_1 and Q_2 are the minimum and maximum values of the input signal voltage E_1 , $p_1(E_1)$ is the probability density function of the input voltage, and $F'(E_1)$ is the slope of $F(E_1)$, the compression characteristic.

Some experimental results obtained with a laboratory model of a quantizer are given in Figs. 6-9. Figs. 6-7 show measurements on the third harmonic associated with 6-digit quantizing. As mentioned before, the amplitude of any one harmonic oscillates with load. The calculated curves shown were obtained by straightforward Fourier analysis. In the measurements it was convenient to spot only the successive nulls and peaks.

In Fig. 6 the bias was set to correspond to the stair-case curve of Fig. 1, while in Fig. 7 the origin is moved to the point $(E_0/2, E_0/2)$, i.e., to the middle of a riser instead of a tread. The peaks of ratio of harmonic to fundamental decrease steadily as the amplitude of the signal is increased to full load, which is just opposite to the usual behavior of a communication system. It is difficult to extrapolate experience with other systems to specify quality in terms of this type of harmonic distortion.

Figure 8 shows measurements of the total distortion power falling in the signal band when the signal is itself a flat band of thermal noise. The technique of making such measurements has been described in earlier articles.^{9,10} Measurements are shown for quantizing with both equal and tapered steps. The particular taper used is indicated by the expander characteristic of Fig. 9. The compression curve is found by interchanging

horizontal and vertical scales. The measurements were made on a quantizer with 32, 64, and 128 steps, and a sampling rate of 8,000 cycles per sec-

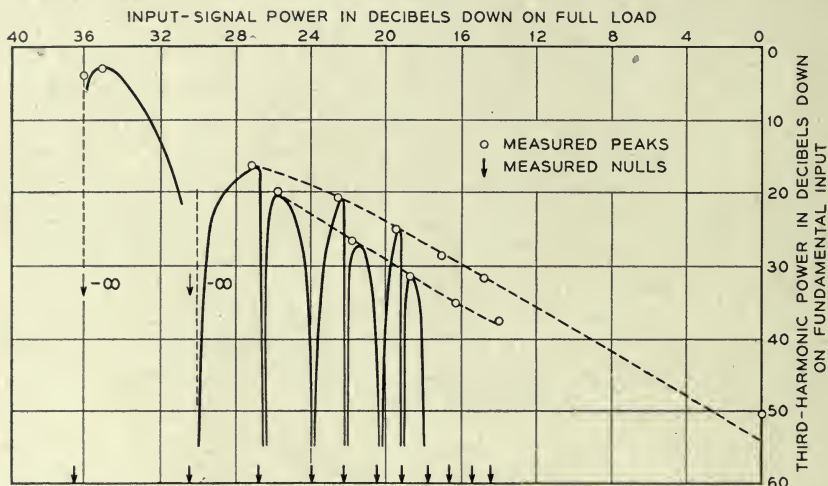


Fig. 6—Third harmonic in 64-step quantized output with bias at mid-tread. The smooth curves represent computed values.

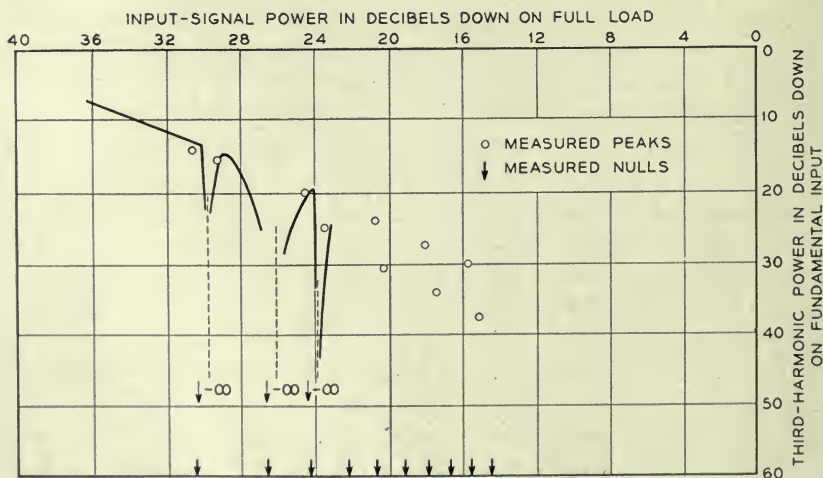


Fig. 7—Third harmonic in 64-step quantized output with bias at mid-riser. The smooth curves represent computed values.

ond. The applied signal was confined to a range below 4,000 cycles per second. With equal steps the distortion power is practically independent of load as shown by the db-for-db straight lines. With tapered steps, the distortion is less for weak signals, and only slightly greater for large signals.

The vertical line designated "full load random noise input" represents the value of noise signal power at which peaks begin to exceed the quantizing

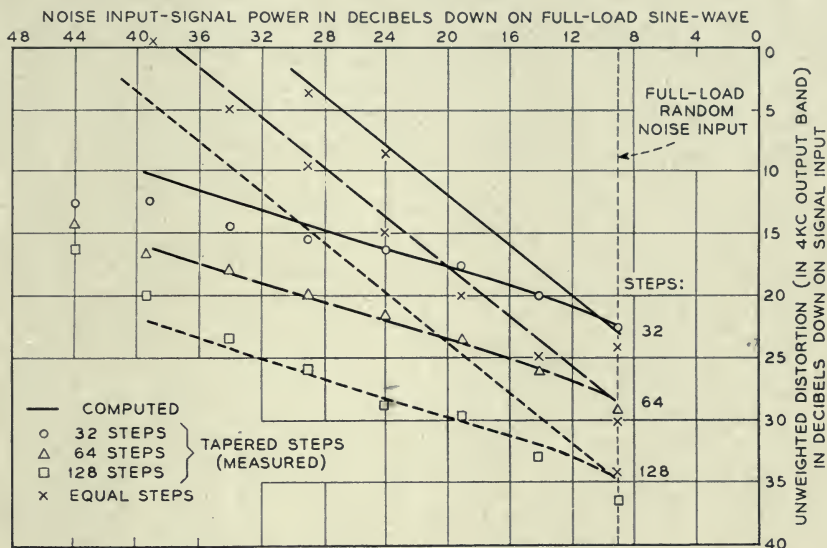


Fig. 8—Total distortion in signal band from quantizing with equal and tapered steps.

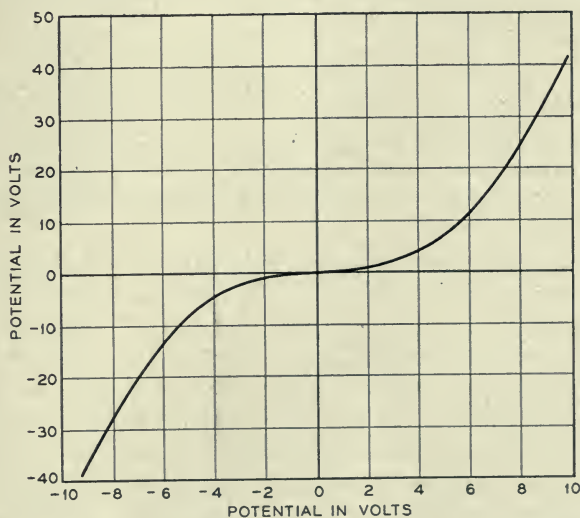


Fig. 9—Expanding characteristic applied to noise in tapered steps of Fig. (8).

range. This occurs when the rms value of input is 9 db below the rms value of the sine wave which fully loads the quantizer.

Flatness of the distortion spectrum with frequency within the signal band is demonstrated by Fig. 10. Two kinds of input were used here—a flat band of thermal noise and a set of 16 sine waves with frequencies distributed throughout the band. Results in the two cases were practically the same. The theoretical levels of distortion power for the band widths of the measuring filters (95 cps) are shown by the horizontal lines.

In the experimental results given here use has been made of laboratory studies by Messrs. A. E. Johanson, W. A. Klute, and L. A. Meacham.

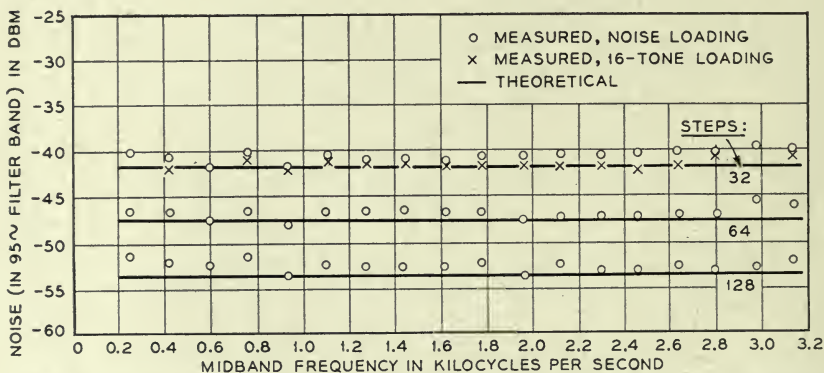


Fig. 10—Spectral density of distortion in signal band from quantizing and sampling. The quantizing steps were equal and the quantizer was fully loaded by a random noise or 16-tone input signal with mean power = -2.5 dbm.

2. THEORETICAL ANALYSIS

The correlation theorem discovered by N. Wiener¹¹ may be stated as follows: Let ψ_τ represent the average value of the product $I(t)I(t + \tau)$, where $I(t)$ is the value of a variable such as current or voltage at time t , and $I(t + \tau)$ is the value at a time τ seconds later. Mathematically:

$$\psi_\tau = \overline{I(t)I(t + \tau)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t)I(t + \tau) dt \quad (2.0)$$

From analogy with statistical theory, ψ_τ is called the correlation of $I(t)$ with itself, or the autocorrelation function of the signal. Since we shall not deal here with the correlation of two signals, we shall shorten our terms and call ψ_τ simply the correlation of $I(t)$. Let $w_f df$ represent the mean power in the output of an ideal bandpass filter of width df centered at f . We assume that the ideal filter is designed to work between resistances of one ohm each and that the input signal $I(t)$ is delivered to the filter from a source with internal resistance of one ohm. (The use of unit resistances does not restrict the generality of the results, since equivalent transmission performance

of any linear electrical circuit is obtained by multiplying all impedances by a constant factor. All voltages are multiplied and all currents divided by the same factor. By assuming unit values of resistance we are able to use

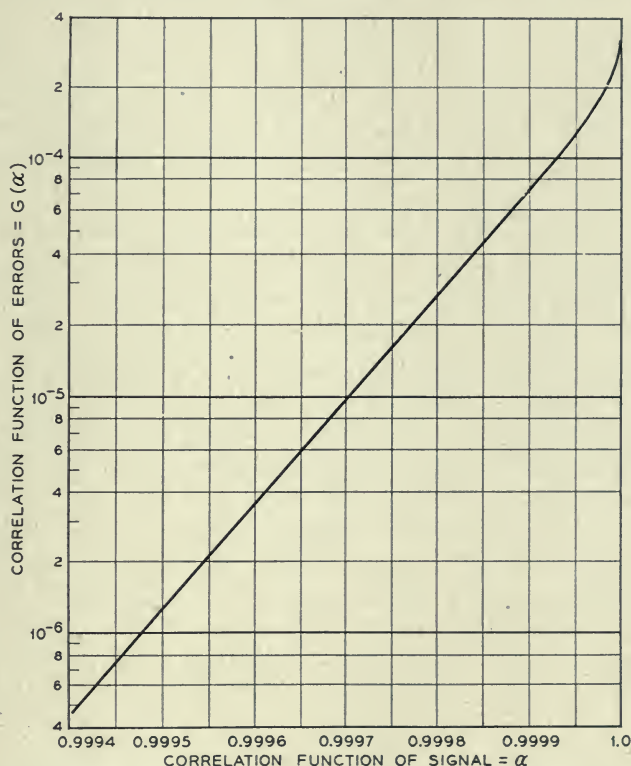


Fig. 11—Correlation function of 7-digit quantizing errors.

squared values of voltages and currents to represent power.) The theorem states that w_f and ψ are related by the equation:

$$w_f = 4 \int_0^{\infty} \psi_{\tau} \cos 2\pi f\tau \, d\tau \quad (2.1)$$

Proof may be found in the references cited. When the signal contains periodic components, the integral in (2.1) becomes divergent in the ordinary or Riemann sense, but this difficulty may be overcome by either applying the theory of divergent integrals or replacing Riemann by Stieltjes integration. We shall not require these modifications here because we shall base our analysis on signals with a continuous spectrum. We note that ψ_0 is the mean

square value of the signal itself. We also point out that the inversion formula for the Fourier integral enables us to express ψ_τ in terms of w_f , thus:

$$\psi_\tau = \int_0^\infty w_f \cos 2\pi\tau f df \quad (2.2)$$

It also may be shown that the ratio ψ_τ/ψ_0 cannot have values outside the interval from -1 to $+1$.

The correlation theorem furnishes a powerful analytical tool for the solution of modulation problems because the calculation of the average ψ_τ is often a straightforward process, while direct calculation of w_f may be a very devious one. Once ψ_τ has been obtained, Eq. (2.1) brings the highly developed theory of Fourier integrals to bear on the computation of w_f .

We shall give the derivation of w_f for quantizing noise making use of the correlation function. In the analysis we shall apply a number of other needed theorems with appropriate references given for proof.

Our first problem is that of calculating the spectrum of the output of the staircase transducer, Fig. 1, when the spectrum of the input signal is given. Let w_f represent the power spectrum of the input signal and ψ_τ the auto-correlation function. The two quantities are related by (2.1) and it is sufficient to express our results in terms of either one. If the instantaneous value of the input signal is represented by E_1 , and that of the output by E_2 , the staircase function may be defined mathematically by:

$$E_2 = mE_0, \quad \frac{2m-1}{2} E_0 < E_1 < \frac{2m+1}{2} E_0, \quad (2.3)$$

$$m = 0, \pm 1, \pm 2, \dots$$

The error is the difference between E_1 and E_2 and may be written as

$$\epsilon(t) = E_1 - E_2 = E_1 - mE_0, \quad \frac{2m-1}{2} E_0 < E_1 < \frac{2m+1}{2} E_0 \quad (2.4)$$

The error characteristic is plotted in Fig. 3.

One approach depends on a knowledge of the probability density function $p(V_1, V_2)$ of the variables $V_1 = E_1$ at time t and $V_2 = E_2$ at time $t + \tau$. The definition of this function is that $p(V_1, V_2) dV_1 dV_2$ is the probability that V_1 and V_2 lie in a rectangle of dimensions dV_1 and dV_2 centered on the point V_1, V_2 of the $V_1 V_2$ -plane. The function $p(V_1, V_2)$ has been calculated for certain types of signals and in theory could be computed for any signal by standard methods. If it is assumed known, we may determine the

correlation function of the error. Let

$$F(V_1, V_2) = \epsilon(t)\epsilon(t + \tau) = (V_1 - mE_0)(V_2 - nE_0),$$

$$\frac{2m-1}{2}E_0 < V_1 < \frac{2m+1}{2}E_0, \quad \frac{2n-1}{2}E_0 < V_2 < \frac{2n+1}{2}E_0, \quad (2.5)$$

$$m, n = 0, \pm 1, \pm 2, \dots$$

Eq. (2.5) defines $F(V_1, V_2)$ as a definite constant value in each square of width E_0 in the V_1V_2 -plane. By elementary statistical theory, the correlation function ξ_τ of the error wave is now

$$\xi_\tau = \overline{F(V_1, V_2)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(V_1, V_2) p(V_1, V_2) dV_1 dV_2 \quad (2.6)$$

The correlation may therefore be calculated since F and p are known functions. The power spectrum Ω_f of the error wave is then equal to the right-hand member of (2.1) with ξ_τ substituted for ψ_τ .

We are interested in the case in which the signal voltage has a smoothly varying spectrum over a specified band. This is a property of a random noise function which has a normal distribution of instantaneous voltages. The two-dimensional probability density function of such a wave is known¹². It is

$$p(V_1, V_2) = \frac{1}{2\pi \sqrt{\psi_0^2 - \psi_\tau^2}} \exp \left[\frac{\psi_0(V_1^2 + V_2^2) - 2\psi_\tau V_1 V_2}{2(\psi_0^2 - \psi_\tau^2)} \right]. \quad (2.7)$$

By inserting this value and that of $F(V_1, V_2)$ from (2.5) in (2.6), making the change of variable:

$$\left. \begin{aligned} V_1 - mE_0 &= E_0 x/2 \\ V_2 - nE_0 &= E_0 y/2 \end{aligned} \right\} \quad (2.8)$$

and adopting the notation,

$$k = E_0^2/\psi_0, \quad \alpha = \psi_\tau/\psi_0, \quad G(\alpha) = \xi_\tau/\psi_0, \quad (2.9)$$

we obtain the following integral determining ξ_τ ,

$$G(\alpha) = \frac{k^2}{32\pi(1 - \alpha^2)^{1/2}} \int_{-1}^1 \int_{-1}^1 xy H(x, y) \exp \frac{-k(x^2 + y^2 - 2\alpha xy)}{8(1 - \alpha^2)} dx dy. \quad (2.10)$$

$$H(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp \frac{-k[m^2 + m(x - \alpha y) + n^2 + n(y - \alpha x) - 2\alpha mn]}{2(1 - \alpha^2)} \quad (2.11)$$

The power density spectrum of the errors is, from (2.1),

$$\begin{aligned}\Omega_f &= 4 \int_0^\infty \xi_\tau \cos 2\pi f\tau \, d\tau \\ &= 4\psi_0 \int_0^\infty G(\alpha) \cos 2\pi f\tau \, d\tau\end{aligned}\quad (2.12)$$

If the signal band is flat from $f = 0$ to $f = f_0$, with no energy outside this band,

$$\alpha = \frac{1}{f_0} \int_0^{f_0} \cos 2\pi f\tau \, df = \frac{\sin 2\pi f_0 \tau}{2\pi f_0 \tau} \quad (2.13)$$

Letting $\gamma = f/f_0$,

$$\Omega_0(\gamma) = \frac{f_0 \Omega_f}{\psi_0} = \frac{2}{\pi} \int_0^\infty G\left(\frac{\sin z}{z}\right) \cos \gamma z \, dz, \quad (2.14)$$

To complete the calculation, we must evaluate the integral (2.10). The first step is to transform the double summation (2.11) into products of single sums by the change of indices:

$$\begin{pmatrix} m+n = m' \\ m-n = n' \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} m = \frac{m' + n'}{2} \\ n = \frac{m' - n'}{2} \end{pmatrix} \quad (2.15)$$

The rearrangement is permissible because the double series is absolutely convergent. The new indices m' and n' also run from minus to plus infinity, but must be either both even or both odd because $m' \pm n'$ is even. On dropping the primes after the substitution is completed, we find

$$\begin{aligned}H(x, y) &= \sum_{m=-\infty}^{\infty} \exp \frac{-k[2m(x+y) + 4m^2]}{4(1+\alpha)} \sum_{n=-\infty}^{\infty} \\ &\quad \cdot \exp \frac{-k[2n(x-y) + 4n^2]}{4(1-\alpha)} + \sum_{m=-\infty}^{\infty} \\ &\quad \cdot \exp \frac{-k[(2m+1)(x+y) + (2m+1)^2]}{4(1+\alpha)} \sum_{n=-\infty}^{\infty} \\ &\quad \cdot \exp \frac{-k[2n+1)(x-y) + (2n+1)^2]}{4(1-\alpha)}\end{aligned}\quad (2.16)$$

A further simplification results from a change of the variables of integration to eliminate the terms in xy . This is done by setting

$$\begin{pmatrix} x = u + v \\ y = u - v \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} u = (x+y)/2 \\ v = (x-y)/2 \end{pmatrix} \quad (2.17)$$

By calculating the Jacobian of the transformation, we find $dx dy = 2 du dv$. The region of integration in the uv -plane is a rhombus bounded by the lines $u \pm v = \pm 1$. We then have:

$$G(\alpha) = \frac{k^2}{16\pi(1-\alpha^2)^{1/2}} \left[\int_{-1}^0 \int_{-1-u}^{1+u} dv + \int_0^1 du \int_{u-1}^{1-u} dv \right] (u^2 - v^2) \\ \exp \left[-\frac{k}{4} \left(\frac{u^2}{1+\alpha} + \frac{v^2}{1-\alpha} \right) \right] \sum_{m=-\infty}^{\infty} \exp \frac{-2mk(2u+2m)}{4(1+\alpha)} \\ \sum_{n=-\infty}^{\infty} \exp \frac{-2nk(2v+2n)}{4(1-\alpha)} + \sum_{m=-\infty}^{\infty} \exp \frac{-(2m+1)k(2u+2m+1)}{4(1+\alpha)} \\ \sum_{n=-\infty}^{\infty} \exp \frac{-(2n+1)k(2v+2n+1)}{4(1-\alpha)} \quad (2.18)$$

If we substitute $u = -x$ in the first double integral, $m = -m'$ in the first series, and $m = -m' - 1$ in the third series, we see that the two double integrals are equal. We therefore drop the first double integral and multiply the second by two. The inner integral may then be split into parts with limits from $v = 0$ to $v = 1 - u$ and $v = u - 1$ to $v = 0$. Substituting $v = -y$ in the second part and treating the series as before, we find that the two parts give equal contributions, so that the bracketed integral terms become

$$4 \int_0^1 du \int_0^{1-u} dv$$

applied to the integrand.

The series in (2.18) may be written as Theta Functions, and the imaginary transformation of Jacobi then used as an aid in reduction. We may proceed in a more direct manner, however, by applying Poisson's Summation Formula:¹³

$$\sum_{n=-\infty}^{\infty} \varphi(2\pi n) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(\tau) e^{-im\tau} d\tau \quad (2.19)$$

We thereby show that

$$\sum_{m=-\infty}^{\infty} \exp [-am(x+2m)] = \\ \sqrt{\frac{\pi}{2a}} e^{ax^2/8} \left[1 + 2 \sum_{m=1}^{\infty} e^{-m^2/\pi^2 2a} \cos \frac{m\pi x}{2} \right] \quad (2.20)$$

$$\sum_{m=-\infty}^{\infty} \exp [-a(2m+1)(x+2m+1)] \\ = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{ax^2/4} \left[1 + 2 \sum_{m=1}^{\infty} (-)^m e^{-m^2\pi^2/4a} \cos \frac{m\pi x}{2} \right] \quad (2.21)$$

When the series in (2.18) of type corresponding to the left-hand members of (2.20) and (2.21) are replaced by the equivalent righthand members, positive exponents containing the squared variables of integration are introduced which cancel the negative exponents already present in the integrand. The resulting integral may be written:

$$G(\alpha) = \frac{k}{4} \int_0^1 du \int_0^{1-u} (u^2 - v^2) [f_1(1 + \alpha, u) f_1(1 - \alpha, v) + f_2(1 + \alpha, u) f_2(1 - \alpha, v)] dv, \quad (2.22)$$

where

$$f_1(a, x) = 1 + 2 \sum_{m=1}^{\infty} \exp \frac{-m^2 \pi^2 a}{2k} \cos \frac{m\pi x}{2} \quad (2.23)$$

$$f_2(a, x) = 1 + 2 \sum_{m=1}^{\infty} (-)^m \exp \frac{-m^2 \pi^2 a}{2k} \cos \frac{m\pi x}{2} \quad (2.24)$$

The integrations may now be performed without difficulty. The complete result, which as we shall immediately show is hardly ever necessary to use in full is:

$$\begin{aligned} G(\alpha) = & \frac{k}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \exp \left(-\frac{4n^2 \pi^2}{k} \right) \sinh \frac{4n^2 \pi^2 \alpha}{k} \\ & + \frac{k}{\pi^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (m \neq n) \frac{1}{(m^2 - n^2)} \exp \frac{-4(m^2 + n^2) \pi^2}{k} \\ & \sinh \frac{4(m^2 - n^2) \pi^2 \alpha}{k} - \frac{k}{\pi^2} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (m \neq n) \frac{1}{(m - \frac{1}{2})^2 - (n - \frac{1}{2})^2} \\ & \exp \frac{-4[(m - \frac{1}{2})^2 + (n - \frac{1}{2})^2] \pi^2}{k} \sinh \frac{4[(m - \frac{1}{2})^2 - (n - \frac{1}{2})^2] \pi^2 \alpha}{k} \quad (2.25) \end{aligned}$$

An alternative derivation of (2.25), subsequently suggested by Mr. S. O. Rice, is based on the fact that $\epsilon(t)$ as defined by (2.4) or Fig. 3 is a periodic function of E_1 which can be expanded in a Fourier series with period E_0 . Substituting the series in (2.5) leads to an expression for $\epsilon(t) \epsilon(t + \tau)$ as the product of two Fourier series. After proof that it is permissible to write this product as a double series and to calculate the average sum as the sum of the averages of the individual terms the problem is reduced to a double series in which the typical term is proportional to the average value of $\exp i(uV_1 + vV_2)$ where u and v are constants depending on the position of the term in the series. Rice has shown¹² that the average value of such a term is $\exp [-(u^2 + v^2)\psi_0/2 - uv\psi_r]$. Summation of these terms leads again to (2.25).

From the defining equation (2.9) we note that k is a small quantity when more than a very few steps are used in the quantizer so that exponentials with exponent containing the factor $-1/k$ are very small except when the factor is multiplied by a number near zero. It will be seen that this can only happen in the first series and then only when α approaches the value unity. We recall that α lies in the range -1 to $+1$ and it is apparent from (2.25) that $G(\alpha)$ is an odd function of α . We thus need consider only positive values of α very slightly less than unity. Only the component of the $\sinh$ with positive exponent is then significant, and we write the very accurate approximation for $G(\alpha)$:

$$G(\alpha) \doteq \frac{k}{2\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \exp \frac{-4n^2\pi^2(1-\alpha)}{k} \quad (2.26)$$

A typical curve of $G(\alpha)$ vs. α for a fixed value of k is shown in Fig. 11. The rapidity with which it falls away at the left of the point $\alpha = 1$ is such that the curve can only be plotted by greatly expanding the scale of α in this region. The physical significance of the spike-shaped curve is that $G(\alpha)$ is a measure of the correlation of the errors as a function of the correlation of the applied signal. When there are many steps there is virtually no correlation between errors in successive samples except when there is complete correlation of successive signal values.

Use of the approximation (2.26) enables us to derive a convenient formula for the spectral density of the errors in a flat band input signal. Substituting (2.26) in (2.14) we obtain:

$$\Omega_0(\gamma) \doteq \frac{k}{\pi^3} \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^{\infty} \exp \left[\frac{-4n^2\pi^2}{k} \left(1 - \frac{\sin z}{z} \right) \right] \cos \gamma z \, dz \quad (2.27)$$

The integrand is negligible except when z is near zero, and in this region we may replace $(\sin z)/z$ by the first two terms of its power series expansion. We then find

$$\begin{aligned} \Omega_0(\gamma) &\doteq \frac{k}{\pi^3} \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^{\infty} \exp \left(\frac{-2n^2\pi^2 z^2}{3k} \right) \cos \gamma z \, dz \\ &= \frac{k}{2\pi^3} \sqrt{\frac{3k}{2\pi}} \sum_{n=1}^{\infty} \frac{1}{n^3} \exp \left(\frac{-3k\gamma^2}{8n^2\pi^2} \right). \end{aligned} \quad (2.28)$$

Only one set of calculations from the infinite series need be made since we may define a function of one variable

$$B(z) = \sum_{n=1}^{\infty} \frac{e^{-z/n^2}}{n^3}. \quad (2.29)$$

Then

$$\Omega_0(\gamma) \doteq \frac{k}{2\pi^3} \sqrt{\frac{3k}{2\pi}} B \left(\frac{3k\gamma^2}{8\pi^2} \right). \quad (2.30)$$

The curves of Fig. (4) were obtained in this way. The relation between k and the number of digits N is based on the assumption of the rms value of signal reaching one-fourth the instantaneous overload voltage of the quantizer. Since zero signal voltage is in the middle of the quantizing range $2^N E_0$, the overload signal measured from zero is $2^{N-1} E_0$. The mean square signal input is ψ_0 . Therefore

$$2^{N-1} E_0 = 4\sqrt{\psi_0} \quad (2.31)$$

or from (2.9)

$$k = 1/4^{N-3} \quad (2.32)$$

We thus have obtained the spectrum of the quantizing errors without sampling. To apply our results to the sampling case we sum up all contributions from each harmonic of the sampling rate beating with the noise spectrum from quantizing only. The resulting power spectrum is given by

$$A_f = \Omega_f + \sum_{n=1}^{\infty} (\Omega_{nf_s-f} + \Omega_{nf_s+f}), \quad 0 \leq f \leq f_s/2. \quad (2.33)$$

If y is the ratio of sampling frequency to signal band width and $A_0(y)$ is the ratio of quantizing power received in the signal band to the applied signal power,

$$A_0(y) = \Omega_0(1) + \sum_{n=1}^{\infty} [\Omega_0(ny + 1) + \Omega_0(ny - 1)]. \quad (2.34)$$

This is the equation used in calculating the curves of Fig. (5).

APPENDIX I

RELATION BETWEEN MEAN SQUARES OF SIGNAL AND ITS SAMPLES

We have already shown that there is a unique relationship between a signal occupying the band of all frequencies less than f_c , and the sampled values of the signal taken at a rate $f_s = 2f_c$. If we are given the signal wave, we can obviously determine the samples; and if we are given the samples, we can determine the signal wave since it is the response of an ideal low-pass filter of cutoff frequency f_c to unit impulses multiplied by the samples. If we apply samples of a signal containing components of frequency greater than f_c , the output of the filter is a new signal with frequencies confined to the band from zero to f_c and yielding the same sampled values as the original wideband signal.

We now consider the problem of determining the mean square value of the samples of an arbitrary function $f(t)$. Let the samples be taken at $t = nT$, $n = 0, \pm 1, \pm 2, \dots$, where $T = 1/2f_c = 1/f_s$.

We may write an expression for the squared samples as a limit of the product of the squared signal and a periodic switching function of infinitesimal contact time, thus

$$f^2(nt_0) = \lim_{\tau \rightarrow 0} f^2(t)S(\tau, t) \quad (\text{I-1})$$

where:

$$S(\tau, t) = \begin{pmatrix} 1, & -\tau/2 < t < \tau/2 \\ 0, & \tau/2 < t < T - \tau/2 \end{pmatrix} \quad (\text{I-2})$$

$$S(\tau, t + T) = S(\tau, t), n = 0, \pm 1, \pm 2, \dots \quad (\text{I-3})$$

By straightforward Fourier series expansion:

$$S(\tau, t) = \frac{\tau}{T} + \sum_{m=1}^{\infty} \frac{2 \sin m\pi\tau/T}{m\pi} \cos 2m\pi f_s t. \quad (\text{I-4})$$

The mean square value of the samples is the limit of the average value of f^2S taken over the contact intervals of duration τ . The average value of f^2S taken over all time, including the blank intervals, is in the limit a fraction τ/T of the average over the contact intervals only. Therefore

$$\begin{aligned} \overline{f^2(nt_0)} &= \lim_{\tau \rightarrow 0} \frac{T}{\tau} \overline{f^2(t)S(\tau, t)} \\ &= \lim_{\tau \rightarrow 0} \overline{f^2(t) + \sum_{m=1}^{\infty} \frac{2T \sin m\pi\tau/T}{m\pi\tau} f^2(t) \cos 2m\pi f_s t} \\ &= \overline{f^2(t)} + \lim_{\tau \rightarrow 0} \sum_{m=1}^{\infty} \frac{2T \sin m\pi\tau/T}{m\pi\tau} \overline{f^2(t) \cos 2m\pi f_s t}. \end{aligned} \quad (\text{I-5})$$

Now the long time average value of $f^2(t) \cos 2m\pi f_s t$ must vanish unless $f^2(t)$ contains a component of frequency mf_s . This could not happen except where $f(t)$ itself contains a component of frequency $mf_s/2$ or two components f_1 and f_2 such that

$$|f_1 \pm f_2| = mf_s \quad (\text{I-6})$$

When no such relation of dependency exists:

$$\overline{f^2(nt_0)} = \overline{f^2(t)}. \quad (\text{I-7})$$

As pointed out before if $f(t)$ contains no frequencies above f_c , the response of the ideal low-pass filter to the samples is $f(t)$, and $f(nt_0)$ represents the samples of $f(t)$. If $f(t)$ does contain frequencies exceeding f_c , the response of the filter is $\phi(t)$, where $\phi(t)$ is wholly confined to the band 0 to f_c and yields the same samples as $f(t)$, i.e.,

$$\phi(nt_0) = f(nt_0), n = 0, \pm 1, \pm 2, \dots \quad (\text{I-8})$$

Eq. (I—7) applied to $\phi(t)$ gives the result:

$$\overline{\phi^2(n t_0)} = \overline{\phi^2(t)}. \quad (\text{I—9})$$

By combining (I—8) and (I—9), we obtain

$$\overline{f^2(n t_0)} = \overline{\phi^2(t)}. \quad (\text{I—10})$$

APPENDIX II

FUNDAMENTAL THEOREM ON APERTURE EFFECT IN SAMPLING

If we sample the wave $Q \cos qt$ at a rate f_s , and multiply each sample by a short rectangular pulse of unit height and duration τ centered at the sampling instants, we obtain by reference to Eq. (I—4) replacing $2\pi f_s$ by ω_s ,

$$F(t) = Q \cos qt S(\tau, t) = \frac{\tau}{T} Q \cos qt + Q \sum_{m=1}^{\infty} \frac{\sin m\pi\tau/T}{m\pi} [\cos (m\omega_s + q)t + \cos (m\omega_s - q)t]. \quad (\text{II—1})$$

The fact that pulse modulation is similar to the more familiar carrier modulation processes is brought out by this equation; the sampling frequency is in fact the carrier. The writer has found that the method of calculation he published in 1933,¹⁴ in which the signal and carrier frequencies are taken as independent variables, is ideally suited for calculations of pulse-modulated spectra. Artificial and cumbersome devices such as assuming the signal and sampling frequencies to be harmonics of a common frequency are thereby avoided.

A unit impulse $\delta(t)$ has zero duration and unit area; hence we may write:

$$\delta(t) = \lim_{\tau \rightarrow 0} \frac{S(\tau, t)}{\tau}. \quad (\text{II—2})$$

A train of samples in which each sample is multiplied by a unit impulse may therefore be written as

$$\sum_{n=-\infty}^{\infty} Q \cos qt \delta(t - t_n) = \lim_{\tau \rightarrow 0} \left[\frac{Q}{T} \cos qt + Q \sum_{m=1}^{\infty} \frac{\sin m\pi\tau/T}{m\pi\tau} [\cos (m\omega_s + q)t + \cos (m\omega_s - q)t] \right]. \quad (\text{II—3})$$

Suppose we apply the train of waves (II—3) to a linear electrical network which delivers the response $g(t)$ when the input is a unit impulse $\delta(t)$. The steady state admittance of the network is given by¹⁵

$$Y_0(i\omega) = \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt \quad (\text{II—4})$$

and the response of the network to (II-3) is therefore:

$$\begin{aligned}
 I(t) = & \frac{Q}{T} |Y_0(iq)| \cos [qt + ph Y_0(iq)] \\
 & + \frac{Q}{T} \sum_{m=1}^{\infty} (|Y_0(im\omega_s + iq)| \cos [(m\omega_s + q)t \\
 & + ph Y_0(im\omega_s + iq)] + |Y_0(im\omega_s - iq)| \cos [(m\omega_s - q)t \\
 & + ph Y_0(im\omega_s - iq)]).
 \end{aligned} \tag{II-5}$$

But $I(t)$ evidently represents a train of pulses in which the pulse occurring at $t = nT$ is equal to the n th sample multiplied by $g(t - nT)$. We have thus obtained the spectrum of a set of samples in which the pulse representing a unit sample is the generalized wave form $g(t)$. Furthermore if the signal frequency q is less than $\omega_s/2$, an ideal low-pass filter with cutoff at $\omega_s/2$ responds only to the first component of (II-5).

The "aperture effect" or variation of transfer admittance with signal frequency is thus given by

$$Y(iq) = \frac{1}{T} Y_0(iq) = f_s Y_0(iq). \tag{II-6}$$

This is Theorem II. Since the system is linear when the signal frequency does not exceed half the sampling frequency, the principle of superposition may be applied to composite signals. In the case of distortion from quantizing errors the aperture effect applies to the error component delivered by the low-pass output filter. For an imperfect low-pass filter in the output we multiply the aperture admittance function by the actual transfer admittance of the filter.

A theorem equivalent to the above has been derived by a different method in a recent paper¹⁶ published after completion of the above work.

REFERENCES

1. W. R. Bennett, Time Division Multiplex Systems, *Bell Sys. Tech. Jour.*, Vol. 18, pp. 1-31; Jan. 1939.
2. H. S. Black, Pulse Code Modulation, *Bell Lab. Record*, Vol. 25, pp. 265-269; July, 1947.
3. W. M. Goodall, Telephony by Pulse Code Modulation, *Bell Sys. Tech. Jour.*, Vol. 26, pp. 395-409; July, 1947.
4. D. D. Grieg, Pulse Count Modulation System, *Tele-Tech.*, Vol. 6, pp. 48-50, 98; Sept. 1947; also *Elect. Comm.*, Vol. 24, pp. 287-296; Sept. 1947.
5. A. G. Clavier, P. F. Panter, and D. D. Grieg, PCM Distortion Analysis, *Elec. Engg.*, Vol. 66, pp. 1110-1122; Nov. 1947.
6. H. S. Black and J. O. Edson, PCM Equipment, *Elec. Engg.*, Vol. 66, pp. 1123-1125; Nov. 1947.
7. L. A. Meacham and E. Peterson, An Experimental Pulse Code Modulation System of Toll Quality, *Bell Sys. Tech. Jour.*, Vol. 27, pp. 1-43; Jan., 1948.
8. H. Nyquist, Certain Topics in Telegraph Transmission Theory, *A. I. E. E. Trans.*, pp. 617-644; April, 1928.
9. E. Peterson, Gas Tube Noise Generator for Circuit Testing, *Bell Lab. Record*, Vol. 18, pp. 81-83; Nov. 1939.

10. W. R. Bennett, Cross-Modulation in Multichannel Amplifiers, *Bell Sys. Tech. J.*, Vol. 19, pp. 587-610; Oct. 1940.
11. N. Wiener, Generalized Harmonic Analysis, *Acta. Math.*, Vol. 55, pp. 177-258; 1930.
12. S. O. Rice, Mathematical Analysis of Random Noise, *Bell Sys. Tech. Jour.*, Vol. 24, p. 50; Jan. 1945.
13. R. Courant and D. Hilbert, *Methoden der Mathematischen Physik*, Vol. 1, p. 64; Berlin, 1931.
14. W. R. Bennett, New Results in the Calculation of Modulation Products, *Bell Sys. Tech. Jour.* Vol. 12, pp. 228-243; April 1933.
15. G. A. Campbell, The Practical Application of the Fourier Integral, *Bell Sys. Tech. Jour.*, Vol. 7, pp. 639-707; Oct. 1928.
16. S. C. Kleene, Analysis of Lengthening of Modulated Repetitive Pulses, *Proc. I. R. E.*, Vol. 35, pp. 1049-1053; 1947.

Analysis and Performance of Waveguide-Hybrid Rings for Microwaves

By H. T. BUDENBOM

This paper presents an analytical treatment of waveguide hybrid rings for microwaves, considered as re-entrant transmission lines. The resulting lines are transformed into equivalent "T" or "lattice" network sections, and determinantal methods are applied in analyzing these equivalent network assemblies for their transmission properties. Some experimental results obtained from a carefully constructed sample of each of two specific types are given. A satisfactory agreement is obtained between the values predicted by theory and experimental results.

INTRODUCTION

IN A recent paper¹, Mr. W. A. Tyrrell has described two general types of waveguide or waveguide/coaxial structures whose properties include bridge or null balance characteristics analogous to those of the hybrid coil common in voice-frequency communication practice. One type, the hybrid junction, is a particular orthogonal junction of four rectangular waveguides. Certain properties of the hybrid junction, notably its impedance characteristics, have been the subject of a British publication². The present paper presents a method for detailed analysis of the other general structure described by Tyrrell, the hybrid ring. This latter structure is essentially an annular ring or annulus of waveguide, at present usually an integral number of quarter wavelengths in circumference, and fitted with an appropriate number of series or shunt branch taps. In this article, phrases such as "quarter wavelength," etc., describing tap spacing or mean annulus perimeter, refer to wavelength in the guide, not to free space wavelength.

The method of analysis employed herein is essentially to treat the tapped annulus as a re-entrant transmission line. Certain circuit equivalences and quarter wave impedance transformations were used by Tyrrell in his paper to develop, with the aid of the reciprocity theorem, many basic properties of hybrid circles and hybrid junctions. In the present paper "T" or "lattice" equivalents (neglecting dissipation) are developed for each section of the annulus, and the method of determinants is applied.

The hybrid junction (known also as the "magic tee") came into use in the newer radars in the latter part of the war. One of its uses, that of providing

¹ "Hybrid Circuits for Microwaves," W. A. Tyrrell, *Proc. I. R. E.*, November 1947.

² "The Theory and Experimental Behaviour of Right-Angled Junctions in Rectangular-Section Wave Guides," *I. E. E. Jour.*, September 1946, p. 177.

as outputs the sum and the difference of two input voltages*, is shown on Fig. 1. Matching stubs at the crossing, as indicated, are required to reduce standing waves to a reasonable value. The corresponding type of hybrid ring for providing sum and difference outputs is likewise shown, to-

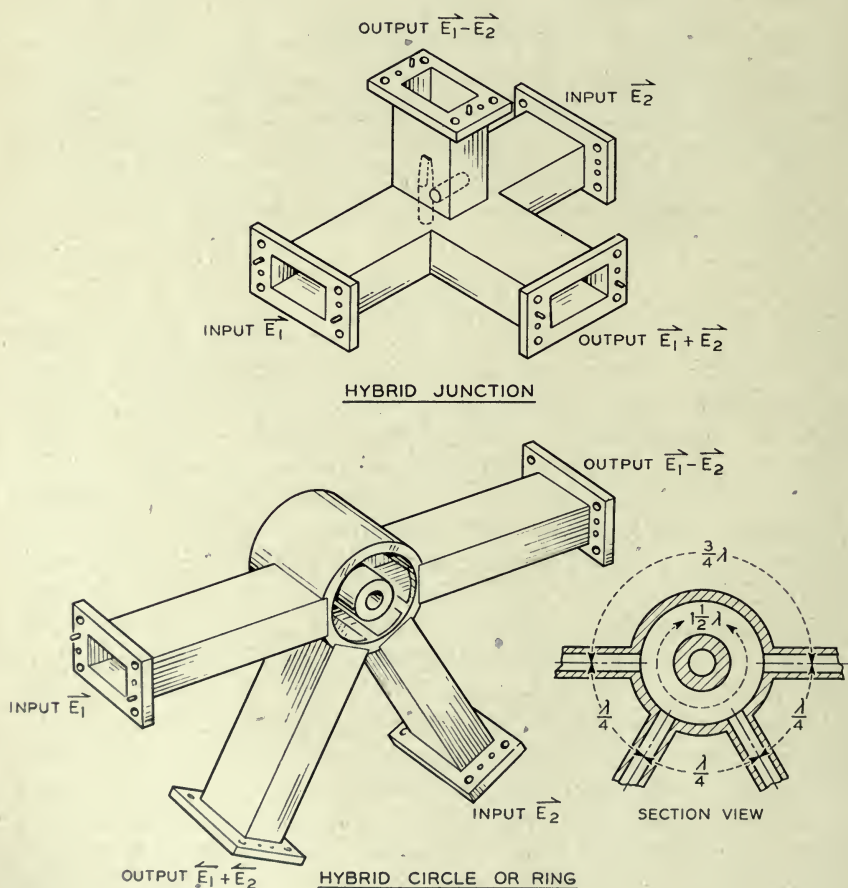


Fig. 1—Hybrid junction and hybrid circle or ring.

gether with a diagram dimensioned in terms of wavelength. Since the path lengths from each input to the output between them are equal, this output gives their sum; the path lengths to the remaining outputs differ by one half wavelength, consequently this output feeds out the difference of the two in-

* More exactly, of two input powers.

puts. No matching stubs are required to achieve a fairly good standing wave ratio; however, the bandwidth over which the ring operates differentially is inherently narrower than that of the junction. The rings have considerably higher power capacity.

The use of hybrids, both junctions and circles, has been noted, as applied to both duplexer and mixer design³.

There follows a circuit analysis of hybrid rings, primarily of the series type. The method used is to consider the annulus as a continuous line closed on itself. The sections between series taps are then treated as being made up of integral single or multiple quarter wave line sections. Equivalent T or lattice sections are derived for 1, 2, 3, and 4 quarter-wavelength sections, ignoring line dissipation. These equivalences are used to draw equivalent mesh networks. The mesh networks are then solved by determinantal methods. To study some effects of frequency shift off the design center, where the mean periphery of the ring departs from an exact integral number of quarter-wavelengths, the increments in the element values for a quarter-wave equivalent T section are calculated and utilized. The example studied is a ring of $1\frac{1}{2}\lambda$ mean perimeter with 3 and 4 taps.

The general procedure neglects possible fringing effects at the junctions. It also neglects the fact that each tap embraces a length of ring which is distinctly more than a small fraction of a wavelength. Nevertheless, the results appear in every case to give a good first approximation. The writer is indebted to Messrs. J. T. Caulfield and J. F. P. Martin for checking the calculations.

Throughout the analysis Z_0 represents guide impedance and $\bar{Z}$ represents annulus impedance. It will be noted that the analytical match condition listed is $\sqrt{2}\bar{Z} = Z_0$ for the $1\frac{1}{2}\lambda$ rings.

The variation of the method necessary to treat the case of shunt taps is indicated.

I. CIRCUIT ANALYSIS

The rings studied herein are of the series type. This type is the one which results when waveguide is bent in the H plane, into a circle, and tap connections are made to the broad outer face. This type of ring is used, for example, in the "rat race" plumbing.

Such rings may be considered on the basis that the annular slot is a transmission line, whose characteristic impedance will here be called $\bar{Z}$ and propa-

³ E. G. Schneider, *Proc. I. R. E.*,—August 1946, p. 528 et seq.—see page 550 et seq. and Figs. 40, 42 and 47.

gation constant P . The transmission line is closed upon itself. Series connections are made by the waveguide connections. The waveguide outlets are assumed, by virtue of their lengths and/or terminations, to present wave-

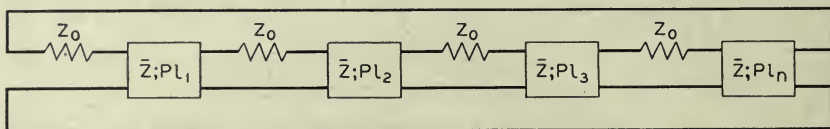


Fig. 2a—Series type hybrid ring as re-entrant transmission line.

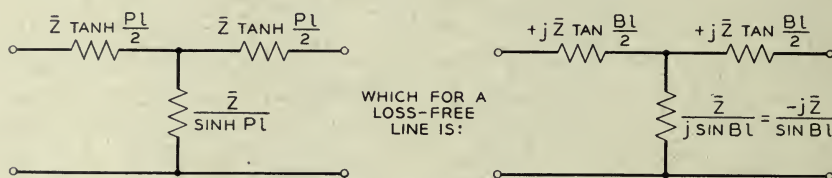


Fig. 2b—T Network equivalent to a line section.

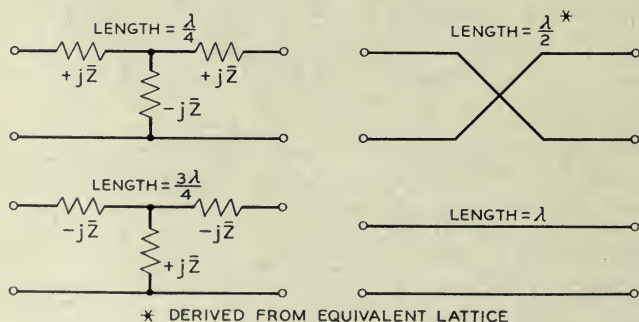


Fig. 2c—Networks equivalent to particular lengths of loss-free line.

* For this case the T becomes indeterminate. However, the needed equivalence can be proved by using the equivalent lattice. If we call Z_a and Z_b the respective series and shunt arms of the T , then the equivalent lattice has series arms $= Z_a$ and diagonal arms $Z_a + 2Z_b$.

guide characteristic impedance to the ring; this will herein be called Z_0 . Diagrammatically, the situation is as in Fig. 2a. In the course of the following, the line sections will be replaced by equivalent networks, assumed non-dissipative.

II. EQUIVALENT LINE SECTIONS

The first method following evaluates the line sections between outlets. The second views each line section as made up of the necessary number of quarter-wave sections, each represented by its equivalent T .

Method 1—The equivalent T for a recurrent structure⁴ of constants $\bar{Z}$ (characteristic impedance) and $P (=A + jB)$ propagation constant per unit length is as shown in Fig. 2b.

There result the equivalences sketched in Fig. 2c.

Once the circuit is diagrammed using the above equivalences, it can be reduced to simpler form by successive combinations of T s, by well known formulae.

Method 2—Determinants. We now consider the line to be made up of the appropriate number of quarter-wave sections, with series taps. Thus we will have Fig. 3.

The shunt impedances are identical; call each Y . The series impedances are made identical by first assuming a tap at each quarter-wave junction;

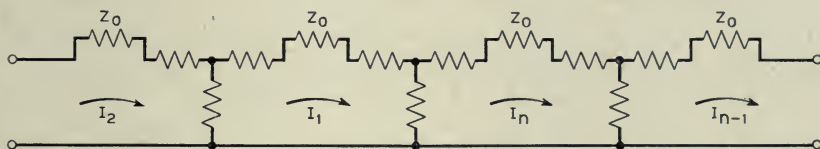


Fig. 3—Re-entrant line as succession of equivalent (quarter wave) T networks and series taps.

call each series leg S . Then the (skew symmetrical) circuit determinant for the case where $N = 10$, (or a $2\frac{1}{2}$ wavelength ring) is

$$D_{10} = \begin{vmatrix} (S+2Y) & -Y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -Y \\ -Y & (S+2Y) & -Y & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -Y & (S+2Y) & -Y & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -Y & (S+2Y) & -Y & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -Y & (S+2Y) & -Y & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -Y & (S+2Y) & -Y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -Y & (S+2Y) & -Y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -Y & (S+2Y) & -Y & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -Y & (S+2Y) & -Y \\ -Y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -Y & (S+2Y) \end{vmatrix} \quad \text{II—} \quad 2.1$$

Now, for the case of an exact integral number of quarter wavelengths around the ring, all $Y_{1-n} = -j\bar{Z}$ and all $S_{1-n} = Z_0 + 2j\bar{Z}$, so all $S + 2Y = Z_0$.

⁴ K. S. Johnson, "Transmission Circuits for Telephone Communication" Book published by D Van Nostrand Co., New York, N. Y.

The determinant then becomes

$$D_{10} = \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +j\bar{Z} & E_1 \\ +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & E_2 \\ 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 & 0 & 0 & 0 & E_3 \\ 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 & 0 & 0 & E_4 \\ 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 & 0 & E_5 \\ 0 & 0 & 0 & 0 & j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 & E_6 \\ 0 & 0 & 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & E_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & E_8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & E_9 \\ +j\bar{Z} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +j\bar{Z} & Z_0 & E_{10} \end{vmatrix} \quad \text{II-2.2}$$

For a $1\frac{1}{2}$ λ ring, $n = 6$ and the system shrinks to

$$D_6 = \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & 0 & +j\bar{Z} \\ +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 \\ 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 \\ 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & +j\bar{Z} & Z_0 \end{vmatrix} \quad \text{II-2.3}$$

For the study of the effects occurring if we move off the design center, we can modify the individual T s to a length $\ell = \frac{\lambda}{4} \pm \frac{\lambda}{N}$. Each series arm, assuming no line dissipation, and N large so $\frac{\lambda}{N} \ll \lambda$, is:

$$\begin{aligned} j\bar{Z} \tan \left[\frac{2\pi}{\lambda} \left(\frac{\ell}{2} \right) \right] &= j\bar{Z} \tan \left[\frac{2\pi}{2\lambda} \left(\frac{\lambda}{4} \pm \frac{\lambda}{N} \right) \right] = j\bar{Z} \tan \left[\frac{\pi}{4} \pm \frac{\pi}{N} \right] \\ &= j\bar{Z} \frac{1 \pm \tan \pi/N}{1 \mp \tan \pi/N} \doteq j\bar{Z}(1 \pm 2\pi/N) = j\bar{Z}(1 \pm \Delta) \end{aligned} \quad \text{II-2.4}$$

Similarly, each shunt arm is:

$$\begin{aligned} \frac{\bar{Z}}{j \sin \left[\frac{2\pi}{\lambda} (\ell) \right]} &= \frac{\bar{Z}}{j \sin \left[\frac{2\pi}{\lambda} \left(\frac{\ell}{4} \pm \frac{\lambda}{N} \right) \right]} = \frac{\bar{Z}}{j \sin \left[\frac{\pi}{2} \pm \frac{2\pi}{N} \right]} \\ &= \frac{-j\bar{Z}}{\sin \frac{\pi}{2} \cos \frac{2\pi}{N}} = \frac{-j\bar{Z}}{\cos \Delta} \doteq \frac{-j\bar{Z}}{1 - \Delta^2/2} \doteq -j\bar{Z} \left(1 + \frac{\Delta^2}{2} \right) \doteq -j\bar{Z} \end{aligned} \quad \text{II-2.5}$$

So the shunt arm is, to a first approximation, not affected by a small shift off design center. Our shunts Y thus remain $-j\bar{Z}$ and

$$S + 2Y = Z_0 + 2j\bar{Z} \pm 2j\Delta\bar{Z} - 2j\bar{Z} = Z_0 \pm 2j\Delta\bar{Z}.$$

Therefore, determinants II—2.2 and II—2.3 can be used by merely considering $\bar{Z}_0 + j2\Delta\bar{Z}$ as a special value of Z_0 .

As is well known,^{5,6} the current solutions are obtained by writing in an external column the driving voltages, opposite their associated meshes, as is done at the right of II—2.2. In the present case the number of driving voltages is usually one, never more than two; so the column will be zeros, save for one (or two) meshes. The current in any mesh is a fraction having D as denominator, and as numerator the minor formed from D by substitut-

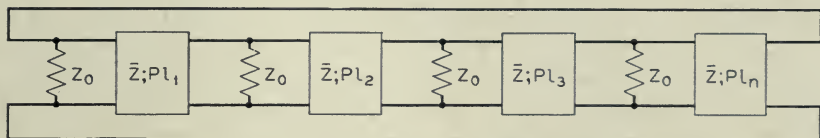


Fig. 4a—Re-entrant line with shunt taps.

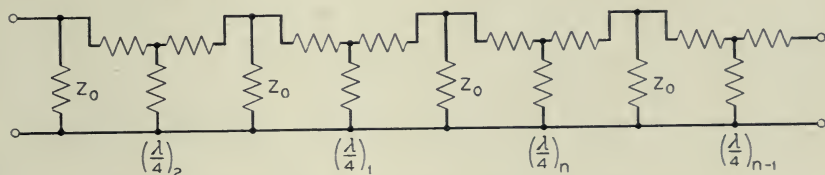


Fig. 4b—Re-entrant line with shunt taps—T networks as line equivalents.

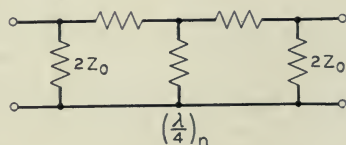


Fig. 4c—Element typical single shunt tapped section.

ing the e.m.f. column in the column corresponding to the mesh where the current is desired, i.e., column n if I_n is desired.

Since D is common to all mesh current expressions, questions of relative power division between branches or of null balance can be handled by operations performed entirely with the numerator minors.

Some slight advantage in evaluating the numerator minors is gained by proceeding where possible so as to make I_1 or I_n the desired current.

⁵ E. A. Guillemin, "Communication Networks," Vols. I and II. Books published by John Wiley and Sons Inc., New York, N. Y.

⁶ L. Silberstein, "Synopsis of Applicable Mathematics." Book published by D. Van Nostrand Co., New York, N. Y.

Alternatively, meshes where $Z_0 = 0$ can be chosen. Another needed quantity is driving point impedance. Since $I_n = \frac{E \cdot d_n}{D}$, then $\frac{1}{Z_{D.P.}} = \frac{d_n}{D}$ and $Z_{D.P.} = \frac{D}{d_n}$. The resulting impedance will include an extra Z_0 , the generator impedance, to which we must match.

It may be of interest to show how the reentrant transmission line analysis can be extended to the case of hybrid rings involving shunt taps. For the reiterative shunt case we have the conditions illustrated in Fig. 4a. With substitution of quarter-wave equivalences Fig. 4a becomes Fig. 4b. Clearly determinants analogous to II—2.1 et seq. can be written for this structure. Alternatively we can split each Z_0 into two parallel impedances, each $2Z_0$, yielding a typical symmetrical section which can be reduced to a simple T or π by well known transformation methods⁴ as shown in Fig. 4c.

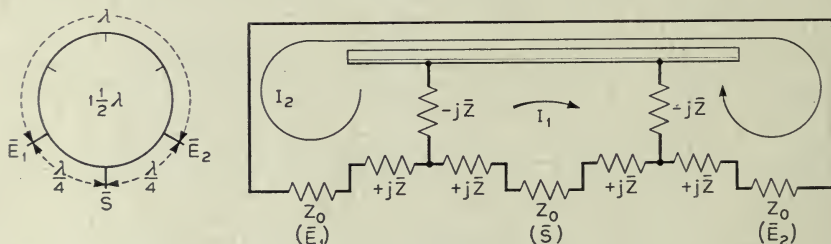


Fig. 5— $1\frac{1}{2} \lambda$ ring—3 arm—equivalent mesh circuit.

III. DETAILED ANALYSIS OF SPECIFIC CASES OF SERIES TYPE RINGS

Case A. $1\frac{1}{2} \lambda$ Ring—3 Arm—As Power Divider—Two Way

This is most simply analyzed using equivalents from Method 1. The equivalent circuit is shown in Fig. 5. It is immediately clear that:

- a. Power fed in at $\bar{S}$ will divide equally between $\bar{E}_1$ and $\bar{E}_2$.
- b. Although $\bar{E}_1$ and $\bar{E}_2$ are in proper wavelength relationship for isolation relative to each other, they are effectively in series and there will not be cancellation. The particular wavelength spacing is thus a necessary but not sufficient condition.
- c. With a voltage E at $(\bar{S})$ we will have

$$E = I_1 Z_0 - I_2 (-2j\bar{Z})$$

$$0 = -I_1 (-2j\bar{Z}) + I_2 (2Z_0)$$

$$\begin{vmatrix} E & Z_0 & +2j\bar{Z} \\ 0 & +2j\bar{Z} & 2Z_0 \end{vmatrix}$$

or

III—1

⁴ loc. cit. page 282.

so

$$\frac{I_1}{E} = \frac{2Z_0}{2Z_0^2 + 4\bar{Z}^2}$$

and the mesh impedance at S is

$$\frac{Z_0^2 + 2\bar{Z}^2}{Z_0} = Z_0 + \frac{2\bar{Z}^2}{Z_0}.$$

Therefore an impedance match is secured if

$$\sqrt{2} \bar{Z} = Z_0. \quad \text{III-2}$$

d. For a voltage e at $\bar{E}_1$, the current at $\bar{E}_2$ may be obtained from

$$\begin{vmatrix} e & 2Z_0 & -2j\bar{Z} \\ 0 & -2j\bar{Z} & Z_0 \end{vmatrix}$$

and is

$$\frac{eZ_0}{2Z_0^2 + 4\bar{Z}^2}. \quad \text{III-3}$$

Under the impedance match condition $\sqrt{2}\bar{Z} = Z_0$, this is $e/4Z_0$ which is just half the current which could be drawn through a load Z_0 connected to a source Z_0 with internal voltage e . Therefore, the "loss" from $\bar{E}_1$ to $\bar{E}_2$ is 6 db.

Case B. $1\frac{1}{2} \lambda$ Ring—4 Arms—As Power Divider and Null Device

As power divider—Two Way (Fig. 6). Using the determinantal method, let $\bar{E}_1$ be in mesh 1. Then $\bar{D}$ is in mesh 4, $\bar{E}_2$ in mesh 5 and $\bar{S}$ in mesh 6. $Z_0 = 0$ for meshes 2 and 3. Then the determinant of II-2.3 and its minor for mesh 5 ($\bar{E}_2$ in Fig. 6) with voltage applied at $\bar{E}_1$, are respectively, from II-2.3:

$$D'_6 = \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & 0 & j\bar{Z} \\ +j\bar{Z} & 0 & +j\bar{Z} & 0 & 0 & 0 \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} & 0 \\ 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & +j\bar{Z} & Z_0 \end{vmatrix} \quad \text{III-3}$$

and

$$d'_6 = \begin{vmatrix} +j\bar{Z} & 0 & +j\bar{Z} & 0 & 0 \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 \\ 0 & 0 & +j\bar{Z} & Z_0 & 0 \\ 0 & 0 & 0 & +j\bar{Z} & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & +j\bar{Z} & Z_0 \end{vmatrix} \quad \text{III-4}$$

Upon expansion d'_6 is found to be 0. Therefore, by adding outlet $\bar{D}$, we have isolated branch $\bar{E}_1$ from branch $\bar{E}_2$. (Compare with case A.) Since

it is well known for this structure that $\bar{D}$ is isolated from input at $\bar{S}$, it must follow that an input at S will still divide equally between $\bar{E}_1$ and $\bar{E}_2$.

As Null Device ($1\frac{1}{2}\lambda$ —4 Arms). We now associate $\bar{S}$ with mesh 1, $\bar{E}_1$ with mesh 2, $\bar{D}$ with mesh 5, $\bar{E}_2$ with mesh 6 (see Fig. 7) which leads to:

$$D_6 = \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & 0 & +j\bar{Z} \\ +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 & 0 \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 \\ 0 & 0 & 0 & +j\bar{Z} & Z_0 & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & +j\bar{Z} & Z_0 \end{vmatrix} \quad \text{III—3.5}$$

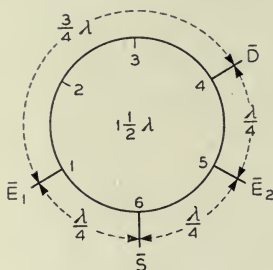


Fig. 6

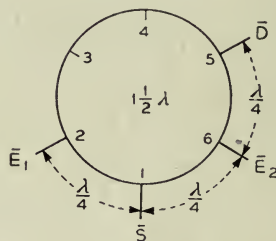


Fig. 7

Fig. 6— $1\frac{1}{2}\lambda$ ring—4 arm—tap spacing and identification for power division analysis by determinants.

Fig. 7— $1\frac{1}{2}\lambda$ ring—4 arm—tap spacing and identification for determinantal analysis as null device.

With voltage applied at $\bar{S}$, mesh 1, the minor for current at $\bar{D}$, mesh 5 is:

$$d_{5(\bar{S})} = + \begin{vmatrix} +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 \\ 0 & 0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & 0 & +j\bar{Z} & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & Z_0 \end{vmatrix} \quad \text{III—5.1}$$

where the $5(\bar{S})$ indicates that the voltage is at $\bar{S}$ and the current is sought at mesh 5. Corresponding minors for the current in mesh $5(\bar{D})$, due to voltages at $\bar{E}_1$ and $\bar{E}_2$ are:

$$d_{5(\bar{E}_1)} = - \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & +j\bar{Z} \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 \\ 0 & 0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & 0 & +j\bar{Z} & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & Z_0 \end{vmatrix} \quad \text{III—5.2}$$

$$d_{5(\bar{E}_2)} = - \begin{vmatrix} Z_0 & +j\bar{Z} & 0 & 0 & +j\bar{Z} \\ +j\bar{Z} & Z_0 & +j\bar{Z} & 0 & 0 \\ 0 & +j\bar{Z} & 0 & +j\bar{Z} & 0 \\ 0 & 0 & +j\bar{Z} & 0 & 0 \\ 0 & 0 & 0 & +j\bar{Z} & +j\bar{Z} \end{vmatrix} \quad \text{III—5.3}$$

Evaluating, we find $d_{5(\bar{S})} = 0$, showing $\bar{D}$ is isolated from $\bar{S}$. The other two expansions give

$$d_{5(\bar{E}_1)} = (+j\bar{Z})[2\bar{Z}^4 - \bar{Z}^2 Z_0^2]$$

and

$$d_{5(\bar{E}_2)} = (+j\bar{Z})[\bar{Z}^2 Z_0^2 - 2\bar{Z}^4]. \quad \text{III—5.4}$$

So the difference between the voltages at $\bar{E}_1$ and $\bar{E}_2$ is transmitted to $\bar{D}$. Hereafter we will operate on a single voltage at $\bar{S}$. Note, incidentally, that if the $\bar{S}$ arm were not terminated the Z_0 in column 1, row 1 of $d_{5(\bar{E}_1)}$ and $d_{5(\bar{E}_2)}$ would be zero, in which case

$$d_{5(\bar{E}_1)} = - (+j\bar{Z})(2\bar{Z}^4) \text{ and } d_{5(\bar{E}_2)} = + (+j\bar{Z})(2\bar{Z}^4) \quad \text{III—5.5}$$

To study frequency shift on current from $\bar{S}$ at $\bar{D}$ we can write III—5.1 as

$$d_{5(\Delta)} = \begin{vmatrix} +j\bar{Z} & Z_0 + 2j\Delta\bar{Z} & +j\bar{Z} & 0 & 0 \\ 0 & +j\bar{Z} & 2j\Delta\bar{Z} & +j\bar{Z} & 0 \\ 0 & 0 & +j\bar{Z} & 2j\Delta\bar{Z} & 0 \\ 0 & 0 & 0 & +j\bar{Z} & +j\bar{Z} \\ +j\bar{Z} & 0 & 0 & 0 & Z_0 + 2j\Delta\bar{Z} \end{vmatrix} \quad \text{III—6}$$

$$= 2\bar{Z}^2(-\bar{Z}^2 j\Delta\bar{Z} + 2Z_0\Delta\bar{Z}^2 + 4j\Delta\bar{Z}^3) = -2\bar{Z}^4 \cdot \Delta\bar{Z}. \quad \text{III—7.1}$$

Match Condition. The impedance match condition is readily shown to be $\sqrt{2}\bar{Z} = Z_0$ as for the three-arm $1\frac{1}{2}\lambda$ ring.

CONSTRUCTION OF TEST SAMPLES

From the drawing of the hybrid circle (Fig. 1), it will be seen that the multiple soldering of guides into the ring can present difficulty in fabrication, especially where numerous branches are required. In addition, early measurements indicated the necessity of accurate dimensions, both linear and angular. As a consequence, the experimental hybrid circles which were used in the measurements reported herein were milled from brass cylinders. Figure 8 shows a 4-branch ring opened so that interior detail can be seen. This form of experimental construction enables dimensions to be held to average values of about half a thousandth of an inch and ten minutes of arc. The mating surfaces are flat to within this tolerance. However, no currents resulting from the field tend to flow in the direction crossing the gap and no loss ensues from this source. These mechanical tolerances are essential only to a basic experiment of the nature here described; larger tolerances could undoubtedly be specified in practice.

EXPERIMENTAL RESULTS

There follows a tabulation of some experimental data on samples of the specific series types analyzed. The attenuation figures are probably good to $\pm .25$ db up to 10 db, to $\pm .5$ db up to 50 db. The SWR figures may not be better than $\pm .2$ db.

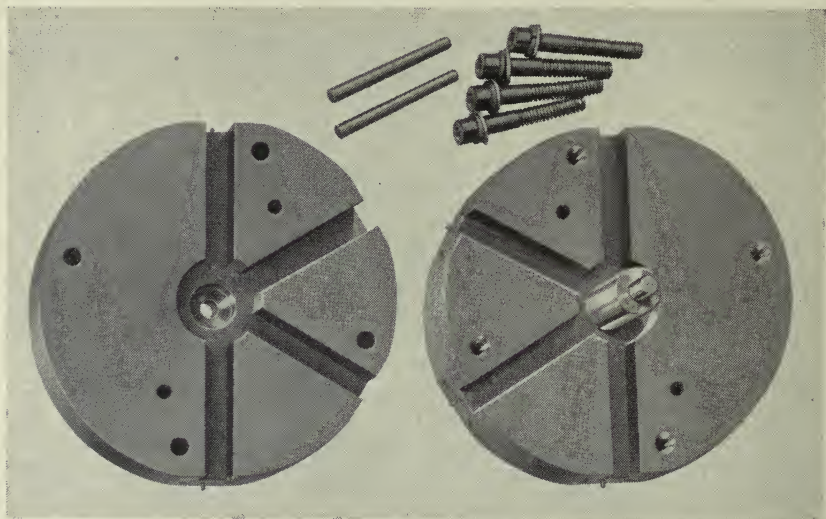


Fig. 8— $1\frac{1}{2}\lambda$ ring—4 arm—photograph of machined test sample.

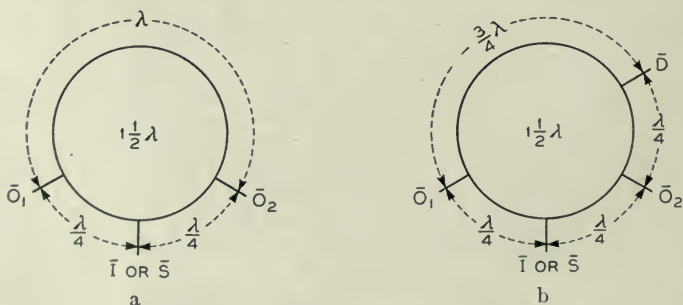


Fig. 9a— $1\frac{1}{2}\lambda$ ring—3 arm—as power divider.

Fig. 9b— $1\frac{1}{2}\lambda$ ring—4 arm—as power divider and null device.

The data are for structures built in terms of .900 inch by .400 inch rectangular guide size (inside) and the test wavelengths* are in the 3-centimeter region. The design wavelength* is λ_0 , the test wavelength λ^* .

Case A: $1\frac{1}{2}\lambda$; Three Arms; Impedance Match $\sqrt{2Z} = Z_0$; Reference Fig. 9a:

* These are space wavelengths.

	Experimental values at $\%(\lambda - \lambda_0)/\lambda_0$				
	-6%	-3%	0	+3%	+6%
Power Division. Input at $\bar{I}$. Relative power output at $\bar{O}_1$ or $\bar{O}_2$ in db (approx. constant over band)	$\longleftarrow$ $\longleftarrow$		-3.7 ($\bar{O}_1$) -3.5 ($\bar{O}_2$)	$\longrightarrow$ $\longrightarrow$	
Transmission Loss (Isolation) between $\bar{O}_1$ and $\bar{O}_2$ in db (approx. constant over band). $\bar{I}$ terminated.	$\longleftarrow$		6.0	$\longrightarrow$	
Standing Wave Ratio (SWR) in db at $\bar{I}$. Outlets $\bar{O}_1$ and $\bar{O}_2$ terminated.	1.10	2.32	.84	1.50	1.20

Case B: $1\frac{1}{2}\lambda$; Four Arms; Impedance Match $\sqrt{2}\bar{Z} = Z_0$; Reference Fig. 9b:

	Experimental values at $100(\lambda - \lambda_0)/\lambda_0$				
	-6%	-3%	0	+3%	+6%
Power Division. Input at $\bar{I}$. Relative power output at $\bar{O}_1$ or $\bar{O}_2$ in db (approx. constant over band).	$\longleftarrow$ $\longleftarrow$		-3.5 ($\bar{O}_1$) -3.5 ($\bar{O}_2$)	$\longrightarrow$ $\longrightarrow$	
Transmission Loss (Isolation) between $\bar{O}_1$ and $\bar{O}_2$ in db $\bar{I}$ and $\bar{D}$ terminated.	20.3		48.5		19.7
Transmission Loss (Isolation) between $\bar{S}$ and $\bar{D}$ in db $\bar{O}_1$ and $\bar{O}_2$ terminated.	24.0		47.7		22.2
Standing Wave Ratio (SWR) in db at $\bar{I}$ ($\bar{S}$). Outlets at $\bar{O}_1$, $\bar{O}_2$ and $\bar{D}$ terminated.	3.50	1.20	.66	.77	2.20

COMPARISON BETWEEN THEORY AND EXPERIMENT

From the experimental results, we can now cite in support of the theory the following areas of agreement between theory and experiment, at the design wavelength:

RING TYPE AND PROPERTY	THEORY	EXPERIMENT
Case A: $1\frac{1}{2}\lambda$; Three Arms		
Relative power at $\bar{O}_1$ and $\bar{O}_2$ for input at $\bar{I}$.	-3 db	-3.6 db
Impedance match (SWR)	0 db	.84 db
Observed center wavelength versus mean annulus perimeter guide wavelength	Agreement to	about 1%
Transmission loss (Isolation) from $\bar{O}_1$ to $\bar{O}_2$. $\bar{I}$ terminated.	6.0 db	6.0 db
Case B: $1\frac{1}{2}\lambda$; Four Arms		
Relative power at $\bar{O}_1$ and $\bar{O}_2$ for input at $\bar{I}$	-3 db	-3.5 db
Impedance Match (SWR)	0 db	.66 db
Transmission Loss (Isolation) $\bar{S}$ to $\bar{D}$. $\bar{O}_1$ and $\bar{O}_2$ terminated	Conjugacy	47.7 db
Transmission Loss (Isolation) $\bar{O}_1$ to $\bar{O}_2$. $\bar{D}$ and $\bar{I}$ terminated	Conjugacy	48.5 db

CONCLUSION

It is concluded that the theory developed provides calculated results in satisfactory accord with experiment.

It will be recalled that the approximation was initially made that the line sections were loss free. The theory could doubtless be extended to include dissipation by retaining a small real component in the propagation constant P of Fig. 2b. No doubt this real component could, in turn, be included to adequate accuracy in the equivalences of Fig. 2c by the addition of real components in the series arms only. That is, the series arm for a $\lambda/4$ section would be $\bar{Z}(r + j1) = r\bar{Z} + j\bar{Z}$ where $r \ll 1$. Since such terms appear as part of the $(S + 2Y)$'s in the basic determinant II—1, which is the same as in series with the Z_0 's in determinant II—2, the inclusion of dissipation would appear to be formally straightforward.

Methods of Electromagnetic Field Analysis*

By S. A. SCHELKUNOFF

This paper presents a discussion of ideas involved in various mathematical methods of electromagnetic field analysis and of the inter-relations between these ideas. It stresses the points of contact between circuit and field theories and their mutually complementary character. While the field theory focuses our attention on the electromagnetic state as a function of position in space, the generalized circuit theory is preoccupied with the electromagnetic state as a function of time. The points of contact between the field and circuit theories are many. Thus, Maxwell's equations are identical with Kirchhoff's equations (really Lagrange-Maxwell equations) of certain three-dimensional networks in which only the adjacent meshes are coupled. The integral equations for the electrical current in conductors embedded in dielectric media are also Kirchhoff equations of certain networks containing infinitely many meshes with a coupling between every two meshes.

From the point of view of electrical performance the difference between a physical network of lumped elements and a continuous network, such as a resonator, is due to a certain difference in the distribution of the zeros and poles of associated impedance functions in the complex impedance plane. Similarly, the difference between ordinary transmission lines and wave guides is due to a difference in the distribution of natural propagation constants.

The paper ends with a general discussion of the discontinuities in wave guides, idealized boundary conditions for simplification of electromagnetic problems, and the analytical character of field vectors regarded as functions of the complex oscillation constant.

IN THE last few years engineering applications of electromagnetic field theory have been greatly expanded. Field theory has become essential for the solution of many practical problems and in planning engineering experiments. New applications have influenced the theory itself and have led to new conceptions. The chasm between the circuit theory of low frequency electrical phenomena and the field theory of high-frequency phenomena has disappeared. The two theories have met in wave guides and their merger has become essential. This paper is a discussion of the essential ideas underlying various mathematical methods of analysis of electromagnetic oscillations and waves in the light of new applications and of the merger of the originally distinct circuit and field theories.

CIRCUIT THEORY

Circuit theory is a mathematical method and it should not be confused with circuits. Empty space is neither a circuit nor a network; but as we shall soon see, for the purposes of analysis the empty space can be treated as a network. It is perfectly true that until recently circuit theory was con-

* This paper was originally delivered as a lecture at a meeting sponsored by the Basic Science Group of the American Institute of Electrical Engineers, April 12, 1945.

cerned almost exclusively with aggregates of "circuit elements" interconnected in various ways. It is also true that the most familiar form of circuit equations is that which is similar to Kirchhoff's equations for the steady current flow in networks of conducting rods, published¹ in April 1845.

This form is applicable only to circuits. However, the application of these "Kirchhoff equations" to alternating currents, natural as it may seem to us now, was not obvious one hundred years ago. The first equation for a simple circuit consisting of a capacitor, an inductor, and a resistor in series was published in 1853 by Lord Kelvin.² Interestingly enough his approach is based on the ideas applicable both to conventional circuits and to high-frequency resonators. If q is the electric charge on one plate of the capacitor, the energy stored in the capacitor is $q^2/2C$, where the coefficient C depends on the geometry of the capacitor. The magnetic energy of the circuit is $\frac{1}{2} L\dot{q}^2$, where $\dot{q}$ is the time rate of change of the charge, that is, the current in the circuit, and L is a coefficient depending on the geometry of the circuit. The rate of energy transformation into heat is $R\dot{q}^2$, where R is a coefficient depending on the geometry of the conductors (and of course on their resistivity). The law of conservation of energy demands that

$$\frac{d}{dt} [q^2/2C + \frac{1}{2} L\dot{q}^2] = -R\dot{q}^2. \quad (1)$$

When the differentiation is performed and $\dot{q}$ is cancelled, the usual form of the equation is obtained. The coefficients of proportionality, that is, the inductance L , the capacitance C , and the resistance R sum up and stress the really important electrical characteristics of the circuit; the details of the construction of the circuit are suppressed.

It was Maxwell who formulated the general equations for electric networks by extending the application of a method developed by Lagrange for mechanical systems. This Maxwell did in his last two lectures. In the words of his student, J. H. Fleming:³ "Maxwell, by a process of extraordinary ingenuity, extended this reasoning (the method of Lagrange) from materio-motive forces, masses, velocities and kinetic energies of gross matter to the electromotive forces, quantities, currents, and electrokinetic energies of electrical matter, and in so doing obtained a similar equation of great generality for attacking electrical problems."

Before discussing the Lagrange-Maxwell method more completely, let us see if we can construct a network whose electrical properties would be the same as those of a continuous medium.

¹ *Annalen der Physik*.

² *Philosophical Magazine*.

³ *Philosophical Magazine*, 1885.

NATURAL NETWORK MODELS OF CONTINUOUS MEDIA AND
MAXWELL'S DIFFERENTIAL EQUATIONS

Transmission line theory represents a well known example of the application of circuit theory to continuous systems. Two-wire transmission lines are subdivided into infinitesimal sections by planes perpendicular to the lines. Each section is replaced by a capacitor whose capacitance is so chosen that, for a given voltage across the transmission line, the electric charges on the plates of the capacitor are correspondingly equal to the charges on the sections of the wires constituting the line. The leads connecting the terminals of these capacitors are then assumed to possess an inductance and a resistance but no capacitance. Thus the electric flux or displacement is "swept" into tiny capacitors, and the magnetic flux or displacement into tiny inductors.

This representation is good only at low frequencies because it depends on the assumption that the electric displacement is only in one direction, namely at right angles to the transmission line. In effect, this representation neglects the capacitance between different parts of the same conductor and includes only the capacitance between the opposite segments of different conductors. That is, while we have recognized that the inductance and capacitance are distributed in the direction parallel to the transmission line, we have ignored the fact that they are also distributed at right angles to the line. In the general representation we should subdivide the medium into infinitesimal blocks and devise a three-dimensional network lattice of infinitely small meshes, Fig. 1. The displacement current can be swept equally into tiny capacitors. If the medium is dissipative, the resistors may be inserted in parallel with the capacitors to take care of the conduction currents in the medium. The magnetic flux is swept equally into tiny coils in the corners of each mesh. However, the resulting network is not homogeneous. Besides meshes of type A consisting of four capacitors and four inductors, it contains meshes of type B consisting of inductors only; and yet we started with a homogeneous medium. Gabriel Kron solved the difficulty by introducing ideal transformers (with one-to-one turn ratio) with their windings in series with the coils at the opposite corners of each A-mesh. These transformers do not affect the electrical performance of the A-meshes but introduce infinite impedance into B-meshes and thus effectively eliminate them.

As a matter of fact, such transformers should properly be included in the network representations of two-wire lines. In fact, by implication they are included as soon as we state that the direct and return currents in the line are equal and opposite. Without an infinite impedance to currents flowing in the same direction we cannot have the balance. Pursuing the matter

further, we should say that all this is in accord with physical facts. The inductance per unit length of an *infinitely long* isolated wire is infinite. The mutual inductance between two parallel wires is also infinite. The two wires are the "windings" of an ideal transformer and a finite impedance is presented only to equal and opposite currents. In the case of wires of finite length the essentially three-dimensional character of the structure manifests itself, and other modes of propagation have to be considered.

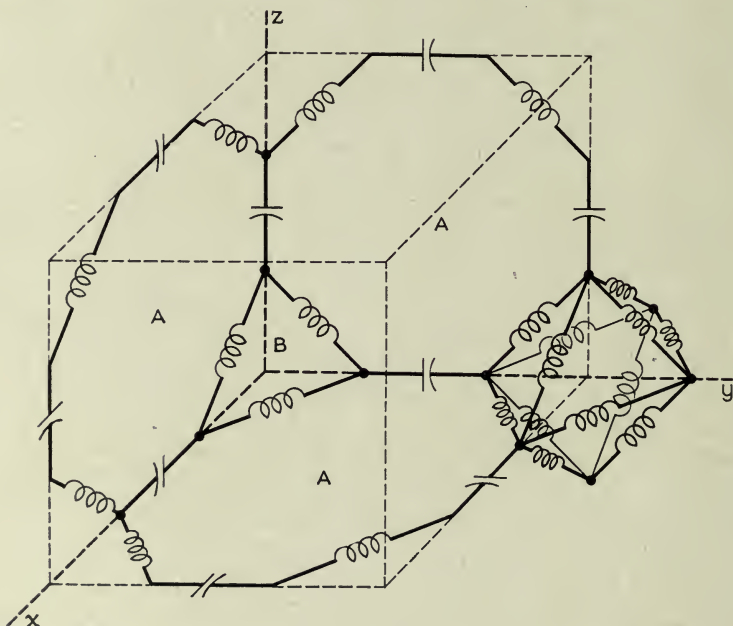


Fig. 1—Typical equivalent meshes in a circuit representation of continuous media.

It is evident that the homogeneity of the medium is not a prerequisite for the existence of its network model. Having the values of L and C at our disposal, we can choose them to reflect the dependence of the permeability μ and the dielectric constant ϵ on position.

If we divide the medium into small blocks of volume $\Delta x \Delta y \Delta z$, the capacitance C_x of the typical capacitor in those branches of the network which are parallel to the x -axis is $C_x = \epsilon \Delta y \Delta z / \Delta x$, where ϵ is the dielectric constant. The conductance in parallel with this is $G_x = g \Delta y \Delta z / \Delta x$. The inductance of the typical coil in the xy -plane is $L_{xy} = \mu \Delta x \Delta y / 4 \Delta z$. The voltages across the capacitors are $E_x \Delta x$, $E_y \Delta y$, $E_z \Delta z$, where E_x , E_y , E_z are the electric intensities, that is, the voltages per unit length in the respective directions. The currents in the coils situated in the xy -plane are equal to $H_z \Delta z$; simi-

larly the currents in the other coils are $H_x \Delta x$ and $H_y \Delta y$. It is to be noted that the capacitors are associated with the corresponding longitudinal components of the electric field while the inductors go with the transverse components of the magnetic field. Applying Kirchhoff's laws to the network in Fig. 1, we should and do obtain Maxwell's field equations. Similarly, we can construct network lattices in the patterns of other coordinate systems, cylindrical and spherical, for example.

Among the obvious conclusions to be drawn from this analysis of the network structure of the medium supporting the electromagnetic field is the validity of certain general network theorems such as the Reciprocity Theorem and Thevenin's Theorem.

REDUCED NETWORK MODELS AND INTEGRAL EQUATIONS OF LORENTZ TYPE

So far we have been concerned with the electromagnetic field in its entirety. In order to visualize the medium as a three-dimensional network we have selected the most direct course: We have subdivided the medium into blocks of displacement current, compressed them into capacitors, and eliminated displacement currents from the rest of space; similarly, we have

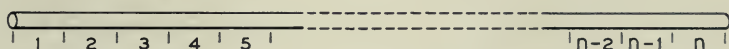


Fig. 2—Subdivision of a straight antenna for its representation by a reduced network with n meshes.

swept the magnetic flux into neat little packages. But this is not the only course open to us. We can suppress the medium just as completely as we normally do in the analysis of elementary networks. In order to illustrate this method let us consider a doublet antenna, Fig. 2. We shall divide it into n sections. The current and charge in any one section exert forces on the charge in any other section. We can regard each section of the antenna as a mesh of a network in which every mesh is coupled to every other mesh. In each mesh the voltage which is necessary to compensate for the electromotive force of self-induction of the mesh itself, for the resistance of the mesh (or rather for the internal impedance of the wire), and for the voltages induced from all the other meshes, is the impressed voltage. The equations assume the following form:

$$\begin{aligned} Z_{11}I_1 + Z_{12}I_2 + Z_{13}I_3 + \cdots + Z_{1n}I_n &= V_1, \\ Z_{21}I_1 + Z_{22}I_2 + Z_{23}I_3 + \cdots + Z_{2n}I_n &= V_2, \\ Z_{n1}I_1 + Z_{n2}I_2 + Z_{n3}I_3 + \cdots + Z_{nn}I_n &= V_n, \end{aligned} \quad (2)$$

where the I 's are the currents in the various sections of the antenna and the V 's are the *impressed* voltages. The Z 's are the self-impedances and the mutual impedances, and are calculated from the law of force between two charged particles. In a transmitting antenna the impressed voltage is zero everywhere except in a restricted region. In the receiving antenna the voltage is impressed on all sections; but one section, the "load," has a very different self-impedance from the remaining sections.

When n is finite, our equations are approximate. If we make n infinite and introduce the impressed electric intensity, that is, the impressed voltage per unit length, we convert equations (2) into a single integral equation. More generally we may have to consider the transverse dimensions of the antenna and divide the entire surface of the antenna into elementary surface elements, each of which will represent *two* meshes in our network. We have to have two meshes for each surface element because the current may in general change its direction from point to point and in order to specify it completely we must consider two components of the current. These may be taken as tangential to some Gaussian coordinate lines drawn on the surface of the antenna. The exact network equations will appear as a system of two integral equations involving double integrals.

In this discussion, we have assumed that the medium outside the antenna is homogeneous. No difficulty is presented by the simultaneous inclusion of a transmitting and a receiving antenna. The two form just one network and the voltages impressed on the various meshes of the receiving antenna represent simply the coupling between these meshes and the meshes of the transmitting antenna. All the mutual impedances are calculable from the general equation,

$$E = -\mu \frac{\partial A}{\partial t} - \text{grad } V, \quad (3)$$

representing the force per unit charge due to a given moving charge. If we so desire, we can take equation (3) with the explicit expressions for A and V in terms of electric current and charge as the fundamental equations of electromagnetic theory and dispense with Maxwell's differential equations altogether. This course is feasible but inexpedient. Actual applications of this equation turn out to be much too complicated in the great majority of practical problems. It is only when we already know the current and charge distribution that (3) becomes really useful. Thus in the accepted development of electromagnetic theory (3) is subordinated to Maxwell's equations and derived from them.

NORMALIZED NETWORK MODEL AND LAGRANGE-MAXWELL ELECTRODYNAMICAL EQUATIONS

Let us now return to the ideas of Lagrange as applied to electromagnetics. In dynamics the Lagrange equations are formulated in terms of the kinetic energy T expressed as a function of velocities, potential energy U expressed as a function of coordinates, and a dissipation function F expressed as a function of velocities. In network theory T is the magnetic energy expressed as a function of currents, U is the electric energy expressed in terms of charges, and F is the dissipation function in terms of currents. Lagrange-Maxwell equations are then written in the following form

$$\frac{d}{dt} \left[\frac{\partial}{\partial \dot{I}_n} (T - U) \right] - \frac{\partial}{\partial q_n} (T - U) + \frac{\partial F}{\partial \dot{I}_n} = V_n, \quad (4)$$

where I_n is the typical mesh current, q_n is its time integral, and V_n is the *impressed* electromotive force, that is, the electromotive force not accounted for by the magnetic induction and the charges in the network. The various functions in the equation are

$$\begin{aligned} T &= \sum_m \sum_n \frac{1}{2} L_{mn} I_m I_n, & U &= \sum_m \sum_n \frac{q_m q_n}{2C_{mn}}, \\ F &= \sum_m \sum_n \frac{1}{2} R_{mn} I_m I_n, \end{aligned} \quad (5)$$

where L_{mn} is the mutual inductance between two typical meshes (the self-inductance if $m = n$), C_{mn} is the mutual capacitance and R_{mn} is the mutual resistance. The mesh currents are introduced in order to insure that the total current either entering or leaving a typical junction of the network elements is zero. If we perform the differentiations indicated in equation (4), we shall obtain the network equations in their usual form.

Let us now suppose that $F = 0$ and $V_n = 0$. In higher algebra it is shown that by a linear transformation two quadratic functions, T and U for example, can be reduced to normal forms in which there are no mutual terms

$$T = \sum_n \frac{1}{2} L_n \hat{I}_n^2, \quad U = \sum_n \hat{q}_n^2 / 2C_n. \quad (6)$$

In this case equations (4) will assume the following simple form

$$L_n \frac{d\hat{I}_n}{dt} + \frac{\hat{q}_n}{C_n} = 0. \quad (7)$$

It is as if we had a certain number of isolated single-mesh circuits. Equations (7) represent the *normal modes of oscillation* of the network.

Take the simple case of two identical coupled circuits, Fig. 3. The network equations are

$$L \frac{d^2 I_1}{dt^2} + \frac{I_1}{C} - M \frac{d^2 I_2}{dt^2} = 0, \quad -M \frac{d^2 I_1}{dt^2} + \frac{I_2}{C} + L \frac{d^2 I_2}{dt^2} = 0. \quad (8)$$

It is evident by inspection that there are two possible modes of oscillation. In one mode $I_1 = I_2$ and in the other $I_1 = -I_2$. The natural frequency of the first mode is $\omega_1 = 1/\sqrt{(L - M)C}$ and that of the second mode $\omega_2 = 1/\sqrt{(L + M)C}$. The magnetic energy function is

$$\begin{aligned} T &= \frac{1}{2} L I_1^2 - M I_1 I_2 + \frac{1}{2} L I_2^2 \\ &= \frac{1}{2} (L - M) \left[\frac{I_1 + I_2}{\sqrt{2}} \right]^2 + \frac{1}{2} (L + M) \left[\frac{I_1 - I_2}{\sqrt{2}} \right]^2. \end{aligned} \quad (9)$$

Thus the sum and the difference of the currents in the two meshes oscillate independently.

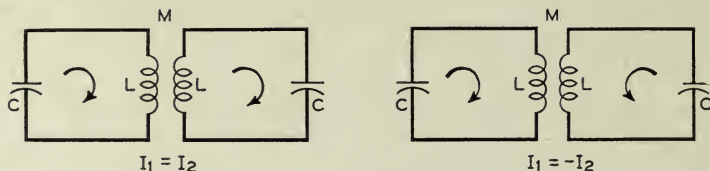


Fig. 3—Two possible modes of oscillation in a symmetric two-mesh circuit.

More generally a network with n meshes possesses n independent modes of oscillation. In each mode the ratios of the mesh currents $I_1, I_2, \dots, I_n$ are prescribed by the network parameters and the connections of the network elements, but the relative strength of the oscillation remains arbitrary. When we pass to networks with distributed parameters such as sections of transmission lines and cavity resonators, we find merely that the number of independent modes of oscillation is infinite. In the case of a nondissipative uniform transmission line with both ends shorted, the natural frequencies of the various oscillation modes are proportional to the sequence of integers: 1, 2, 3, . . . The current distribution for the n -th mode is given by $\sin(n\pi x/\ell)$, where ℓ is the length of the section; but the actual amplitude remains arbitrary. For the gravest mode ($n = 1$) the middle part of the line section behaves as a capacitor and the ends as inductors. For the higher modes the line is subdivided into sections, some of which act primarily as capacitors and others as inductors.

In the case of cavity resonators of some simple shapes, such as parallelepipedal, cylindrical and spherical, the determination of the oscillation modes is a fairly simple problem. The dynamical equations of the resonator

(Maxwell's field equations) are partial differential equations. Their solutions would normally involve arbitrary functions; but since the tangential electric intensity vanishes at the conducting boundary of the resonator, the solutions assume a much less arbitrary form involving only an infinite set of arbitrary constants. Particular solutions are sought in the form of products of three functions, each depending on only one coordinate. For parallelepipedal cavity resonators the various components of electric and magnetic intensity are assumed in the form $X(x) Y(y) Z(z)$. By substituting in Maxwell's equations it is found—very fortunately indeed—that X , Y , Z may be obtained as solutions of ordinary differential equations. The boundary conditions at the boundaries of the box $x = 0, a$; $y = 0, b$; $z = 0, c$ are easy to satisfy because we have to work with only one of these three functions at a time.

In general, however, the problem of calculating oscillation modes is by no means simple; but once these modes have been determined, the problem of forced oscillations as well as free oscillations is practically solved. For instance, a small loop inside a resonator is coupled to the various modes and the coupling coefficients can be determined by evaluating the flux linkages.

Every physical circuit possesses an infinite number of degrees of freedom and circuits with a finite number of degrees of freedom are abstractions. If we take special measures to concentrate magnetic energy as much as possible in a few regions of the medium and electric energy in a few other regions, we shall have a physical network in which a finite number of oscillation modes will be well separated on the frequency scale from all the rest. If we are concerned only with the frequencies comparable to the natural frequencies of this cluster of modes, we can ignore all the higher modes and for our purposes we may regard the network as a finite network. At these frequencies the infinitely small meshes into which we could subdivide the individual "inductors" (regions of magnetic energy concentration) and "capacitors" (regions of electric energy concentration) will oscillate in unison in groups.

Briefly we can summarize the above methods of analysis as follows: The medium supporting the electromagnetic field may be regarded as a three-dimensional network of infinitely small meshes in which every mesh is coupled only to the adjacent mesh. Circuit equations applied to this network lead to Maxwell's differential equations. In contrast with this "*natural network model of the medium*" we can construct a *reduced network model* in which only the conductors of the medium are subdivided into meshes. The medium surrounding the conductors is concealed in the mutual impedances of the constituent meshes. Every mesh is coupled to every other mesh and the mutual impedance (or the coupling factor) is

determined from the law of force exerted by a moving charge on a stationary charge. This approach leads to one or two integral equations which can be approximated by a system of linear algebraic equations. While the latter may seem much simpler than the differential equations obtained from the natural network model, in reality their solution would often constitute a much more difficult analytical problem. The natural network model in which each mesh is coupled only to the adjacent meshes is in harmony with the idea of continuous propagation of electromagnetic disturbances; while the reduced network model conforms to the action at a distance philosophy. The difference is merely in the language and ideas and not in substance.

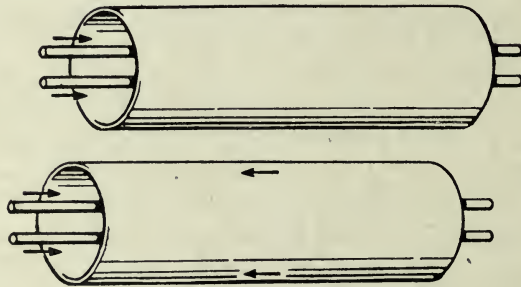


Fig. 4—Two possible modes of propagation in a symmetrically shielded parallel pair.*

Finally, the third method is based on the idea that at certain frequencies, called the natural frequencies, various parts of a closed system oscillate in phase or 180° out of phase, that the most general natural oscillation is the sum of such oscillations, and that the most general forced oscillation can be expressed in terms of fields associated with the natural modes of oscillation. We may call this the *normalized network model of the electromagnetic field*. Thus far we have described it with reference to closed systems or cavity resonators. In effect we have assumed that the amounts of magnetic and electric energy are finite or else we could not talk about T and U functions. The method can be extended to open systems of wave guides.

MODES OF TRANSMISSION

Let us begin with a coaxial transmission line. Everyone is familiar with the particular mode of transmission in which equal and opposite currents flow in the two conductors. The circuit is completed through the dielectric where the displacement current flows from one conductor to the other. Next, consider a shielded parallel pair. If the structure is symmetric, we shall recognize at once two modes of transmission, Fig. 4. In one mode, the balanced mode, the currents in the wires are equal and opposite; there are

* In the upper part of this figure one of the directional arrows should be reversed.

also equal and opposite currents in the shield which, however, are not equal to the corresponding currents in the wires. In the other mode, the currents in the wires are equal and similarly directed, the return path being through the shield; this mode is similar to the coaxial mode since the wires act in parallel, effectively as one conductor. In the case of n wires there are n distinct modes of transmission. Each mode is characterized by the ratio of currents in the wires and by the field pattern that goes with it.

In all these modes the longitudinal current paths are conductive; but there is no reason whatsoever why the circuit closure should not take place through the dielectric. Even in those modes of transmission in which all longitudinal current paths are conductive, we have to depend on the dielectric for completion of the circuit; this should prepare us for the idea that conductors are not essential for wave transmission. If we include the dielectric, the number of possible longitudinal tubes of flow becomes infinite and so does the number of possible transmission modes; but as the cross-section of each individual tube decreases the longitudinal capacitance also decreases, and these modes will participate in the transfer of power over substantial distances only at correspondingly higher frequencies. It is not merely that at low frequencies the longitudinal impedance becomes very high; it is capacitive and causes high attenuation. The effect is analogous to the attenuation in high-pass filters below the cutoff.

The mathematical analysis which lends quantitative substance to these ideas is similar to that involved in the cavity resonator problem. Once all the modes of transmission have been found, the next problem is that of the excitation of these modes by a given source, that is, of coupling of the source to various modes.

To summarize: A physical transmission line or a wave guide has always an infinite number of transmission modes either independent or substantially independent of each other. It is as if we had a system of single-mode transmission lines without couplers. For each transmission mode the structure behaves as a high-pass filter. If n is the number of conductors, there are $n - 1$ transmission modes with the cutoff frequency equal to zero. Since the lowest non-zero cutoff frequency corresponds to a wavelength comparable to the transverse dimensions of the guide, it is clear that in systems with two or more conductors we have a certain finite number of transmission modes which are well separated on the frequency scale from all the rest. For this reason we may ignore all the higher modes when we are concerned with transmission of low frequencies only, by "low" meaning the frequencies well below the frequency equal to the velocity of light divided by the largest transverse dimension of the transmission line.

Analysis of waves in free space proceeds along similar lines. An electric

"dipole" is the source of the simplest spherical electromagnetic wave. We may picture this dipole as a pair of small spheres connected by a thin rod. Under the influence of an impressed force the charge is made to surge back and forth between the spheres. We cannot have a simple source like a uniformly expanding and contracting sphere as in the case of sound waves. The electric charge is conserved, and the only way we can alter the charge in one place is to transfer it to some other place. A more symmetrical dipole would be a single sphere on the surface of which the charge is made to move back and forth between two hemispheres. Let us call these hemispheres respectively the "northern" and the "southern". When the positive charge accumulates on the northern hemisphere, the radial displacement current flows outwards from it. At the same time an equal radial displacement current flows toward the southern hemisphere. The situation is analogous to the balanced mode of transmission along parallel wires, with the two half spaces acting as "the wires". The distance along the line is the distance from the dipole. The radial transmission line is capacitively loaded but the series capacitance increases as the square of the radius and therefore the capacitive series admittance decreases as the reciprocal of the square of the radius. Hence, at some distance from the dipole, the wave propagation will be quite unimpeded just as in ordinary transmission lines free from loading. Near the dipole the series capacitance is high, and the power carried by the wave in comparison with the energy stored is small.

In the next spherical mode of transmission the polar regions of the spherical generator are similarly charged while the opposite charge is concentrated in the equatorial zone. The zonal character of the radial current distribution persists at all distances from the generator. As might be expected the reactive field in the vicinity of a small "tripole" generator is even stronger than in the case of the dipole source.

The sequence of zonal modes of transmission can be continued indefinitely. Next we could imagine tubular modes in which the space surrounding the generator is subdivided into conical tubes with the radial current in adjacent tubes flowing in opposite directions. This picture is essentially physical; but it corresponds very closely to the mathematical expansion of the general solution of Maxwell's equations in spherical harmonics.

FIELD REPRESENTATION IN TERMS OF FIELDS OF SPECIAL TYPES

From the mathematical point of view the method which we have just been considering is based on the idea of representation of the general field in terms of particular fields having certain relatively simple properties. The method is analogous to that employed in circuit theory when the

response to the general electromotive force is expressed in terms of responses to the unit step function, or the unit impulse function, or the steady state responses at various frequencies.

There are numerous variations of the same general idea, some of which are more suitable to one class of problems and others to another class. If the distribution of electric charge and current is known, then in many cases (but not in all) it is best to subdivide it into small volume elements. Except for a possible static electric charge distribution, the elements will be dipoles. The entire field can thus be regarded as the resultant of spherical waves generated by dipoles of given moment and position. To simplify the integration involved in this method certain auxiliary functions, called the retarded potentials, are introduced. One should not try to ascribe to these auxiliary mathematical functions any physical significance and one should always remember that on certain occasions potential functions, other than the retarded potentials, turn out to be more useful. We should also keep in mind that, in order to apply this method, we have to know the complete distribution of electric conduction currents and as a general rule we do not have this information. Consider, for instance, the problem of electromagnetic shielding. The current in the coil is given; but that in the shield has to be determined. There are methods for calculating the induced current; but these methods give at the same time the shielding effectiveness, and that without employing retarded potentials. It is in approximate studies of radiation patterns of antennas and antenna arrays that the retarded potential method is displayed to the best advantage.

The retarded potentials are based on representation of fields in terms of spherical coordinates; that is, in terms of fields associated with hypothetical *point sources* at the origin of the coordinate system. General fields can also be expressed in terms of cylindrical coordinates and, consequently, in terms of fields associated with hypothetical *line* sources situated along the axis of the coordinate system. Likewise, fields can be expressed in cartesian coordinates; that is, in terms of "plane waves". All such representations have useful applications. The current in the coil is given.

DISCONTINUITIES

In the analysis of the various transmission modes for a given wave guide it is assumed at first that the boundaries of the wave guide are analytic functions of the coordinates. Any discontinuity or irregularity has to be treated separately, simply because there is nothing in the analytic part of the wave guide to suggest that a discontinuity might occur, or to prescribe the properties of this discontinuity. Discontinuities may be accidental, unavoidable or intentional. A kink in a wire is an example of an accidental

discontinuity. Open air wire lines have to be supported on poles which, together with the insulators, constitute unavoidable discontinuities. The beginning and the end of a line are always present. Usually these latter discontinuities are simply unavoidable; but, in radio, at least one discontinuity, the antenna, is made to serve a useful purpose. It is clear that the generator and the load connected by a two-wire line, Fig. 5, are dipoles which will generate spherical waves as well as the wave guided by the transmission line. At low frequencies the length of the dipoles is so small compared with the wavelength that the field does not reach out into

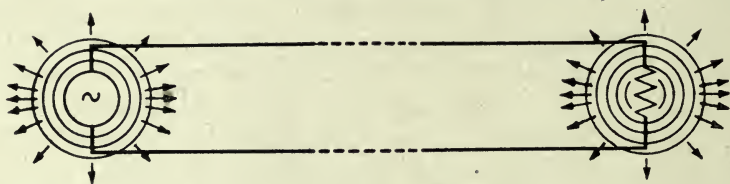


Fig. 5—Formation of spherical waves at the ends of a long pair of parallel wires.

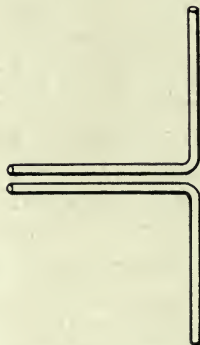


Fig. 6—An antenna.

the region where the radial capacitance becomes negligible and where the spherical wave starts carrying off all the energy that gets there. Spherical waves generated at the beginning and the end of the transmission line are practically stationary waves and constitute merely local reactive reservoirs of energy. The energy is withdrawn from the generator or the transmission line during one half of the cycle only to be returned during the other half. At low frequencies the energy thus exchanged back and forth is so small that normally we don't even think about it. The antenna, Fig. 6, is designed to be a more efficient transformer of the plane wave guided by the parallel pair into the spherical wave which will carry off power to distant points.

Quite frequently discontinuities are introduced intentionally in order to

discriminate against some frequencies. A capacitor in parallel with the wave guide or an inductor in series with it will favor transmission of low frequencies at the expense of high frequencies. These discontinuities are deliberately designed to be sufficiently large to produce noticeable effect. A frequency filter is a more elaborate structure made up of capacitors and inductors designed to achieve desired frequency discrimination.

Discontinuities in high-frequency wave guides are also either accidental, unavoidable or intentional. The principal difference is in the order of magnitude—any irregularity of apparently small physical dimensions may represent a large virtual reservoir of energy. Among the simplest types of intentional discontinuities in wave guides are “irises”, Fig. 7. Local fields are created in the vicinity of the irises. Under the influence of a wave traveling along the guide, electric charge and current are induced in the metal partition. On either side of the partition the complete field is the result of the superposition of fields representing various transmission modes. The cutoff frequencies of these modes may be arranged in an

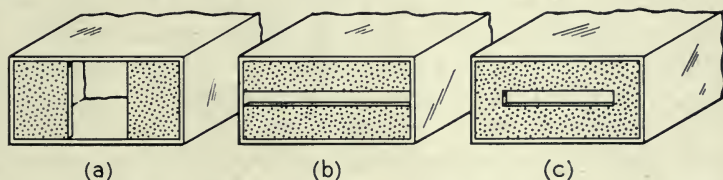


Fig. 7—Inductive, capacitive, and resonant irises.

increasing sequence. If the operating frequency is between the lowest cutoff frequency and the next higher, the propagation constants of all modes except the dominant are real and the corresponding fields will not extend very far from the iris. During one-half cycle the local field withdraws energy from the dominant wave—this being the only source of energy—and during the remaining half this energy is returned. The local field acts as a *virtual source* of power—“virtual” since it operates on borrowed power. On account of symmetry the dominant waves generated by this virtual source and traveling in opposite directions will be of equal intensities. The *scattered wave* traveling toward the source of the incident wave is called *the wave reflected from the iris*; on the other side the scattered and incident waves merge into the *transmitted wave*. The storage of energy in the local field depends on the frequency—hence, the frequency selectivity.

In the case shown in Fig. (7a) the flow of current in the partition is unimpeded and there is no tendency for any local concentration of charge in the partition; the local field is largely magnetic and the iris represents an

inductive reactance. Since any variation of the magnetic field with time always creates an electric field, there will be some capacitance in parallel with the inductance. The same idea may be expressed by saying that the inductance of the iris is not quite independent of the frequency. This lack of constancy is not peculiar to ultra-high frequencies; it is true of coils at low frequencies. Likewise, even at very low frequencies the inductance varies with the frequency because of skin effect.

In the iris shown in Fig. (7b) there are alternating charge concentrations on the upper and lower partitions. The local field is largely electric and the iris is capacitive. A feeble magnetic field associated with charging current is unavoidable, of course; this is also true of capacitors at low frequencies but this time the effect is greater. Finally, an iris of the type shown in Fig. (7c) may be designed to behave as an antiresonant circuit.

In that frequency range in which only the dominant wave is an effective carrier of power to great distances, any discontinuity will behave as a reactive T or Π -network—assuming that observations are made at some distance from the iris where the local field is too feeble to count. This could not be otherwise since there are three parameters at our disposal: two reflection coefficients for waves traveling in opposite directions and one transmission coefficient across the discontinuity. The Reciprocity Theorem requires that the transmission coefficients in the two directions be equal. These three parameters determine the ratios of the reactance elements of the equivalent T or Π -network to the characteristic impedance of the guide.

If the operating frequency exceeds the second cutoff frequency, other waves besides the dominant become effective carriers of power and the equivalent network for the iris becomes more complicated. The iris behaves not only as a dissipative impedance to the dominant wave but also as a negative resistance, to one or more higher order waves.

BOUNDARIES

So far we have paid little attention to the boundaries of the electromagnetic field. Strictly speaking, in any actual situation the field always extends to infinity; the only boundaries there are, are the geometric boundaries between media with different electromagnetic properties. This means that we should solve electromagnetic equations for each homogeneous region, or region with analytically varying properties, and then match the solutions at the boundaries. In many cases, however, this procedure would be very complicated and quite unnecessary. In the case of a cylindrical metal tube with a dipole as a source of power the exact solution may be represented as a particularly formidable integral; but experimentally

we would not be able to detect any difference between the "exact" solution and a much simpler approximate solution.

In the case of rectangular tubes we don't even know how to obtain the "exact" solution in any form; but good approximate solutions are exceedingly simple. The word "exact" is in quotation marks because there can be no really exact solutions of actual physical problems. In the first place the properties of materials are not known exactly; the boundaries between media do not exist in the exact sense of the term; and we just don't know the exact laws of nature. All we really want of any solution is to be accurate enough for some particular purpose. And here is where the idea of idealized boundaries helps in the formulation of simplified, clear-cut mathematical problems. The idea lends flesh and blood to idealized mathematical boundary conditions. *Perfect conductors* have long been mentioned in literature as idealizations of good conductors; but other types of boundaries are of much more recent origin. Perfect conductors are *boundaries of zero surface impedance*; they support electric currents of finite strength when the tangential electric intensity is zero. At these boundaries the tangential magnetic intensity is different from zero. The natural counterpart is a *boundary of infinite impedance* at which the tangential magnetic intensity vanishes but the tangential electric intensity does not. The further generalization is a boundary with a given finite surface impedance which is defined as the ratio of two mutually perpendicular tangential components of the electric and magnetic intensity. The boundary may be isotropic, with its *surface impedance* the same in all directions; likewise, the boundary may be anisotropic. The surface impedance is defined as the ratio of the tangential components of E and H . Since it is necessary to adopt a convention regarding "positive directions" of E and H , these are so chosen that a right-handed screw will advance into the boundary if its handle is turned through 90° from the positive direction of E to coincide with the positive direction of H . In accordance with this convention the positive real part of the surface impedance is associated with an average flow of power into the boundary—that is, with a *passive boundary*. An *active boundary* is a boundary with a negative surface resistance; such boundaries may be used to represent idealized generators of electromagnetic waves and to eliminate from explicit consideration the internal mechanisms of these generators.

FIELD EQUATIONS

Thus far I have tried to present the ideas behind the physical and mathematical analysis of electromagnetic transmission phenomena. These are broader than the electromagnetic laws themselves and, with some super-

ficial modifications, would apply to sound waves, for instance. There are two fundamental equations of transmission of an electromagnetic state, expressing Faraday's law of induction of an electromotive force by a magnetic displacement current and Ampère-Maxwell's law of induction of a magnetomotive force by an electric current. In their most general mathematical form the equations are

$$\oint E_s ds = -\frac{\partial}{\partial t} \iint \mu H_n dS, \quad (10)$$

$$\oint H_s ds = \iint \rho \cdot v_n dS + \iint g E_n dS + \frac{\partial}{\partial t} \iint \epsilon E_n dS,$$

where the subscript s indicates components tangential to a closed path of integration and the subscript n designates components normal to any surface bounded by this closed path. Thus on the left we have "sums" of infinitesimal emf's and mmf's as we travel round some closed curve either on the surface of a wire or just in free space, and on the right we have total magnetic and electric currents linked with this curve. According to our present physical conceptions the magnetic current is always a displacement current defined as the time rate of change of magnetic flux or "displacement". Not that there is anything inconceivable about an actual flow of magnetic charge; it is simply that so far there has been no satisfactory evidence of its existence. In the mathematical analysis it has long been a custom to consider magnetic charges of opposite signs as if they existed; but this is merely for convenience.

The electric current, on the other hand, consists of three components: the *convection current* whose density is the product of the electric charge density ρ and the velocity v ; the *conduction current* whose density is proportional to the electric intensity (the gE term in the above equation) and the *displacement current* defined as the time rate of change of the electric displacement. Strictly speaking, the conduction current is a convection current but of such a kind that it would be extremely awkward to think of it in terms of charged particles and their velocities.

At the same time the statistical result of the irregular movements of these particles can be expressed, for purposes of transmission of an electromagnetic state, as a continuous movement of charge encountering some resistance. There are, of course, such phenomena as resistance noise which are thus automatically excluded from consideration.

In general to these electromagnetic transmission equations we should add the dynamical equations of motion of electric charge; this is essential when dealing with vacuum tubes. But, in considering passive transmission systems, we either omit the convection current altogether, or else assume

that the velocities of the charged particles are specified, and that the forces which they exert on each other are completely neutralized by the forces external to the field, in which case the convection current appears merely as an "impressed current".

Except for the above restrictions, equations (10) form a complete set; but for mathematical convenience two other equations are usually adjoined. These are

$$\begin{aligned}\iint \epsilon E_n dS &= q, \\ \iint \mu H_n dS &= 0,\end{aligned}\tag{11}$$

where the double integration is extended over a closed surface. The first of these equations states that the total electric displacement through a closed surface is equal to the net enclosed electric charge; the second denies the physical existence of magnetic charge. These equations can be derived from (10) and for this reason are not quite on the same footing with them.

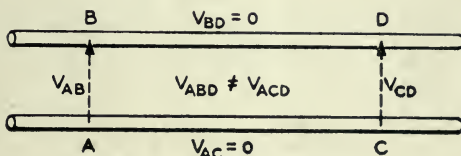


Fig. 8—A pair of parallel wires.

Equation (10) tells us that, except when the field is static, we cannot speak of the electromotive force or the voltage between two points without specifying the path along which we add up the elementary voltages. In fact, equation (10) gives us the difference between the voltages along two different paths connecting the same pair of points. To illustrate, consider a wave along a pair of perfectly conducting wires, Fig. 8. Voltages V_{AC} and V_{BD} along the wires are equal to zero; transverse voltages V_{AB} and V_{CD} are usually unequal; hence $V_{ABD} \neq V_{ACD}$.

If two points are infinitely close, then we can define the voltage unambiguously as the product $E_s ds$ of the electric intensity and the distance between the points. The difference between this voltage and the voltage along any other infinitesimal path is an infinitesimal of the second order, being dependent on the area enclosed by the two paths. In practice two points are sufficiently close if the distance between them is small compared with one quarter wavelength.

Since, except in electrostatics, we cannot speak of the voltage between two points without specifying the path, we cannot speak of the *potential*

difference. In mathematical terms we should say that the differential voltage in a varying electromagnetic field is not an exact differential. To illustrate: $2x \, dx + 2y \, dy$ is an exact differential equal to $d(x^2 + y^2)$ and for this reason its integral depends only on the difference between the values of $(x^2 + y^2)$ at the end points of the path of integration; but $2x \, dx + 2x \, dy$ is not an exact differential and cannot be integrated except when y is given in terms of x so that the path of integration is prescribed.

If equations (10) are applied to infinitesimal closed curves, the following differential equations are obtained:

$$\text{curl } E = -\mu \frac{\partial H}{\partial t}, \quad \text{curl } H = gE + \epsilon \frac{\partial E}{\partial t}. \quad (12)$$

The expressions $\text{curl } E$ and $\text{curl } H$ are merely the symbols for the maximum emf's and mmf's per unit area. These equations are not as general as (10) because they assume that E and H are continuous and at least once differentiable. The equations do not hold across the boundary between different media, where they have to be supplemented by the so-called *boundary conditions* which are obtained from (10). Equations (12) do not hold at a wavefront where E and H are discontinuous; there also we have to supplement them by appropriate boundary conditions, which connect the solutions on the two sides of the wavefront.

ANALYTIC FUNCTIONS

An advance of fundamental importance is made when the field intensities are represented by complex quantities $E e^{j\omega t}$ and $H e^{j\omega t}$ where ω is the frequency in radians. The equations become

$$\text{curl } E = -j\omega\mu H, \quad \text{curl } H = (g + j\omega\epsilon)E, \quad (13)$$

and are thus freed from one independent variable, the time t . This does not mean that we have restricted our analysis to steady state fields; Fourier analysis supplies a general rule for passing from steady states to any state whatsoever. Computational difficulties are great but no greater than they would be in any other method.

A still more important advance is made when the field intensities are represented by $E e^{pt}$, $H e^{pt}$, where the *oscillation constant* $p = \xi + j\omega$ is a complex number. The equations become

$$\text{curl } E = -p\mu H, \quad \text{curl } H = (g + p\epsilon)E. \quad (14)$$

The solutions of these equations are analytic functions of the complex variable p and a way is open for application of the theory of functions of a complex variable.

Thus if we write

$$E = \sum_{n=0}^{\infty} e_n p^n, \quad H = \sum_{n=0}^{\infty} h_n p^n, \quad (15)$$

and substitute in (14), we obtain

$$\begin{aligned} \text{curl } e_0 &= 0, & \text{curl } e_{n+1} &= -\mu h_n, \\ \text{curl } h_0 &= g e_0, & \text{curl } h_{n+1} &= g e_{n+1} + \epsilon e_n. \end{aligned} \quad (16)$$

If these equations are solved subject to the prescribed boundary conditions, E and H will be expressed as power series in the oscillation constant p .

The function theory has already been used successfully in the *restricted circuit theory*; that is, in the theory of finite networks composed of ideal (independent of the frequency) resistances, inductances and capacitances. Likewise, some very general theorems have been established concerning any *physical* input impedance. Whereas the poles and zeros of a function can be anywhere in the complex p -plane, the poles and zeros of the input impedance of a passive system never lie to the right of the imaginary axis. This leads to a theorem to the effect that all poles and zeros on the imaginary axis are simple. The resistance components of the input impedance on the imaginary axis determine the reactance component and hence the complete impedance function except for a purely reactive impedance. The zeros and poles of an impedance occur always in conjugate pairs. These are some of the general theorems of impedance analysis. Not very long ago I came across an expression for the input impedance of a spherical antenna which was obtained by what appeared superficially as a straightforward conventional method; but as soon as I observed that some poles were situated to the right of the imaginary axis, I knew that the expression had to be false. The existence of poles in this region meant a possibility of oscillations which would increase indefinitely of their own accord.

The difference between finite and infinite networks consists in that the former possess a finite number of zeros and poles. All physical structures always possess an infinite number of such singularities; but a finite number of them may form a cluster in the vicinity of the origin, far removed from all other zeros and poles. When this happens we have a physical finite network. In a reactive network all zeros and poles lie on the imaginary axis. In a slightly dissipative system these zeros and poles move a little to the left of the imaginary axis. This happens, for instance, in the case of a thin antenna. The field in the vicinity of a thin wire is large and the radiated power is only a small fraction of the stored energy. The distribution of poles (the solid circles) and zeros (the hollow circles) is illustrated

in Fig. 9. The zero frequency is always a pole for an open type antenna and a zero for a perfectly conducting loop antenna. As the frequency passes through a zero, the antenna impedance passes through a minimum. As the frequency goes through a pole, the antenna impedance passes through a maximum. The disposition of zeros and poles gives us a qualitative idea of the behavior of the impedance as the frequency varies.

As the radius of the antenna increases, the zeros and poles move farther to the left of the imaginary axis. At the same time some zeros and poles, which for a thin antenna are so far to the left that they have very little effect on the impedance, move nearer the origin. For spherical antennas the number of zeros and poles around the origin is considerably larger than for thin doublets.

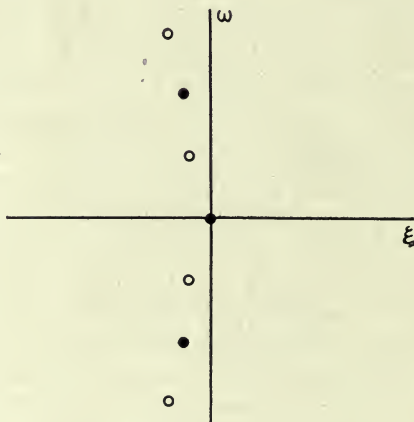


Fig. 9—Distribution of zeros and poles in a dipole antenna: solid circles represent poles; hollow circles zeros.

CIRCUIT AND FIELD EQUATIONS

In conclusion I should like to make a few remarks on the relationship between Kirchhoff's circuit equations and Maxwell's field equations. Are the former approximations; and, if so, in what sense? The answer depends on what is meant by Kirchhoff's equations, for their meaning has changed with passing years. It was exactly a hundred years ago that Kirchhoff stated his equations in a kind of postscript to his paper in *Poggendorf Annalen*; but he contemplated only the d-c networks. Yet nowadays we interpret these equations in such a way that they are applicable to a-c circuits. Some thirty years went by before Maxwell thus generalized the original Kirchhoff equations with the aid of Lagrange's concepts. Maxwell wrote his circuit equations (not the field equations) in a form applicable only to networks with a finite number of degrees of freedom; but nowadays

we interpret these equations in such a way that we can apply them to one-dimensional transmission lines. In so doing we refrain from making approximations which we normally make when applying Kirchhoff laws to networks of lumped elements. In the latter case it is usual to ignore the inductance of the connecting leads or rather the inductance associated with the loop formed by the leads; but in the case of two-wire transmission lines the "connecting leads" constitute the entire network and the loop inductance is no longer ignored. In the case of lumped networks the capacitance between the connecting leads is normally neglected; but this capacitance is scrupulously included in the analysis of two-wire lines since in this case the "lead capacitance" is all the capacitance there is. And I have already referred to a recent contribution of Kron's who presented a three-dimensional network such that if we apply Kirchhoff's laws to it, we shall obtain Maxwell's field equations. The merger between the two points of view is now complete. In its growth, each theory has developed concepts peculiar to itself. The net result is that we are now in a position to understand electromagnetic phenomena better than ever.

The Evolution of the Quartz Crystal Clock*

By WARREN A. MARRISON

SOME of the earliest documents in human history relate to man's interest in timekeeping. This interest arose partly because of his curiosity about the visible world around him, and partly because the art of time measurement became an increasingly important part of living as the need for cooperation between the members of expanding groups increased. There are still in existence devices believed to have been made by the Egyptians six thousand years ago for the purpose of telling time from the stars, and there is good reason to believe that they were in quite general use by the better educated people of that period.¹ Since that period there has been a continuous use and improvement of timekeeping methods and devices, following sometimes quite independent lines, but developing through a long series of new ideas and refinements into the very precise means at our disposal today.

The art of timekeeping and time measurement is of very great value, both from its direct social use in permitting time tables and schedules to be made, and in its relation to other arts and the sciences in which the measurement of rate and duration assume ever increasing importance. The early history of timekeeping was concerned almost entirely with the first of these and for many centuries the chief purpose of timekeeping devices was to provide means for the approximate subdivision of the day, particularly of the daylight hours.

The most obvious events marking the passage of time were the rising and setting of the sun and its continuous apparent motion from east to west through the sky. The first practical measure of the position of the sun of which any record is known was the position or the length of shadows of fixed objects, resulting through a long period of development in the well-known sundial in its many forms. But the sundial was in no sense an instrument of precision and in no sense could be considered as a time *keeping* device. Even after the development which resulted in mounting the gnomon parallel with the axis of the earth, the largest, most elaborate, and most carefully made instruments could at best indicate local solar time. Furthermore, the sundial has value only in daylight hours and then only on

* The subject matter of this paper was given before the British Horological Institute in London on the occasion of the presentation of the Horological Institute's Gold Metal for 1947 to Mr. Marrison in consideration of his contribution toward the development of the quartz crystal clock. The present text is substantially as published in the *Horological Journal*.

days when the sun shines clearly enough to cast a shadow. These shortcomings became more and more important with advances in society and, for measuring duration, man soon began inventing timekeeping means that would work without benefit of the sun.

The evolution of timekeeping devices may be divided into three main periods, each employing a specific type of method, although overlapping to some degree in their applications, and characterized by increasing orders of accuracy.

A graphical representation of this evolution, indicating these three periods of development, and showing the relation between some of the major contributions to time keeping and the resulting accuracy of time measurement, is shown in Fig. 1. The methods employed chiefly during these three periods may be classified broadly as CONTINUOUS FLOW from the beginning up until about 1000 A.D., as APERIODIC CONTROL from then until about 1675 A.D. and as RESONANCE CONTROL from that time up to the present. Keeping in mind the logarithmic nature of the time and accuracy scales used in this graph, it can be seen readily that most of the advancement has been made in a very small part of the total time, corresponding to the resonance control epoch.

THE EPOCH OF CONTINUOUS FLOW

Perhaps due to a feeling that the passage of time was like the flow of some medium, the first time *measuring* devices were those depending on the flow of water into or out of suitable basins. It was recognized that, with an orifice properly chosen, the time required to fill or empty a given basin should be about the same on repetition, and hence was born the first reliable means for measuring time at night or on overcast days. A great variety of devices operating on this principle were constructed and used, some of the earliest having been made by the Babylonians and the Egyptians 3500 years ago.

Some of these water clocks, or clepsydra as they were called, had floats or other indicators which were intended to subdivide a unit of time into substantially uniform divisions. Others were constructed so that successive fillings of the basin would be counted or would operate a stepping device, associated with a dial or other indicator. Through the centuries great numbers of such devices were constructed, with some of the later ones having elaborate mechanisms for striking the hours or for animating figures of people or animals.

For use in places where water was not readily available and where sand was plentiful, clepsydra were developed that would operate with the flow of sand in much the same way as with the flow of water. The basic ideas were not greatly different, the substitution being merely one of expedience.

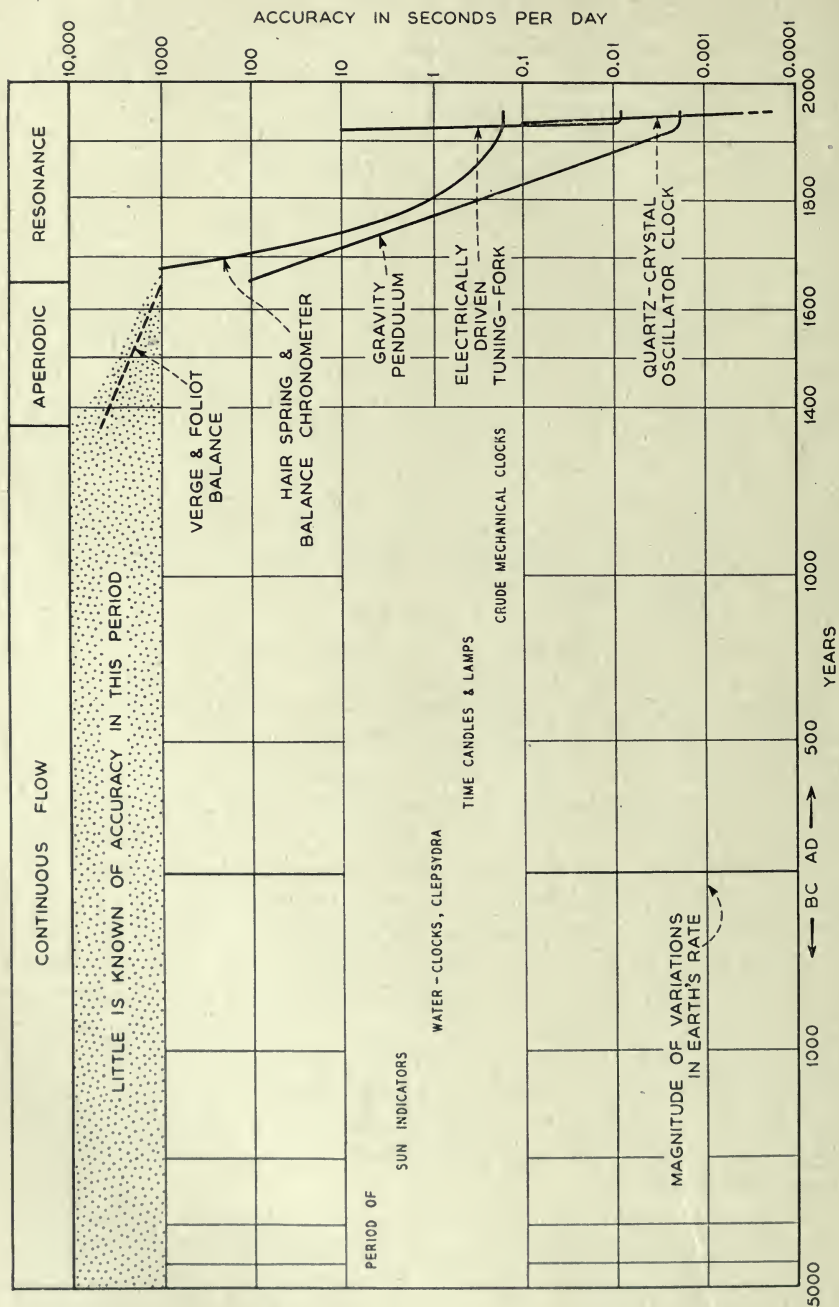


Fig. 1—The accuracy of timekeeping through history.

The hour glass, and its smaller counterparts, is one of the most convenient forms of this device and until quite recent times served a useful purpose where accuracy was of no great importance. The hour glass shown in Fig. 2 was used by a pastor in the early eighteen hundreds to determine the length of his sermons. The average variation among a set of ten one-hour determinations made recently with this glass was 3 minutes, or about 5 per cent.

The clepsydra that were designed to repeat and totalize an endless succession of cycles were especially adaptable to the measurement of extended intervals of time, although with very poor accuracy as we now think of it.



Fig. 2—Hour glass.

By suitable design any desired number of cycles could be made equal to the natural large unit, the day, so that any fraction of a day within the accuracy of a given instrument could be determined simply by counting off the number of cycles from a particular starting point such as sunrise, sunset, or high noon. It was possible with these devices to operate without calibration over periods of several days, although the cumulative error inevitably was very large.

An error of a few hours was of small importance in the days when the speed of communication and travel alike depended on pack animals or the caprices of the wind. And so, in spite of the inaccuracies of the water clocks and sand clocks, they served their purpose well through many centuries.

In fact, it was not until the tenth century A.D. that any really novel effort was made to improve upon them as timekeepers. The first efforts to improve upon them, making use of falling weights for motive power and various frictional devices to control the rate of fall, were not very successful because no satisfactory means were known to keep a friction-controlled device sufficiently constant for the job. Clocks so constructed were no better timekeepers on the whole than the traditional clepsydra. They had, however, the hope of compactness, and much ingenuity was exercised in their design over several centuries.

Also in the category of continuous flow devices should be mentioned the methods depending on the rate of burning, such as in time candles, time lamps and their numerous variations. Such timekeepers are not very accurate but are thoroughly reliable in dry, quiet places, even providing their own illumination at night. Such timekeepers are known to have been used before the tenth century A.D. and certain variations still are used by a few isolated tribes, especially in the tropics.

THE EPOCH OF APERIODIC CONTROL

In or about the year 1360 the invention of an escapement mechanism for controlling an alternating motion from a steady motive power, such as a suspended weight, was the first really important step in the history of precision clock development, and marks the beginning of the second major epoch in timekeeping evolution. The escapement in one form or another was soon applied in practically all timekeepers, the most outstanding example of an early application being a clock constructed by Henry De Vick for Charles V of France in or about the year 1360 A.D. and still in use—with extensive modifications—in the Palais de Justice in Paris.

This invention was important, not because De Vick's clock, or any of its immediate successors, were good timekeepers, but because this was the first time that vibratory motion in a mechanism was used deliberately to control the rate of a time-measuring device. All precision clocks depend in one way or another on using energy to produce vibratory motion, and on using the rate of that motion to regulate suitable dials and other mechanisms.

No simple improvement on De Vick's clock could ever have produced a precision clock in the modern sense, however, because the essential rate-controlling feature was still lacking. His invention consisted of the use of a verge escapement which produced oscillatory motion in a dynamically balanced member, known as a foliot balance, having essentially only moment of inertia and friction. The rate of oscillation, therefore, depended to a large extent on the applied force exerted by the falling weight through a train of wheels, and upon the friction of the escapement parts and of the oscillating member itself.

This sort of operation is known sometimes as relaxation oscillation and appears in many forms. In the clock, the rate-controlling feature depends upon the length of time it takes a member having a given moment of inertia to move from one angular position to another under a given applied torque. Thus, the rate depends to first order on the applied torque.

Although De Vick's clock was one of the most famous in all history, it was not because of its good record of timekeeping. In its original form, it is said that it often varied as much as two hours a day from true time. Outwardly, this clock on the Palais de Justice appears about the same as it did originally, but the "works" have been modernized and it keeps much better time now.

The history of timekeeping during the next three hundred years consisted mainly in improvements and in a great variety of applications of the principles contained in De Vick's clock. During this period great numbers of clocks of all sizes, from tower clocks to portable table clocks were made, controlled by various forms of the crown wheel, verge and foliot balance. All of these timekeepers belong to the class that we have just called aperiodic. Their accuracy, in general, was still poor and the indicator on their dials consisted of but one hand—the hour hand. It was not until the invention and application of the pendulum that the next major improvement was born in timekeeping.

THE EPOCH OF RESONANT CONTROL

All that has been said so far is a prelude to the shortest but by far the most productive epoch in timekeeping, that of resonant control. The heart of every precision clock is an oscillatory device which depends upon *resonance* for its constancy of rate. The history of precision clock development consists largely of the choice and design of stable resonant elements and of devising means for using them so that as far as possible their inherent properties alone control their rates of oscillation. Once in stable oscillation, it is only necessary to control the indicating of dials and other suitable mechanisms in order to constitute a complete clock.* Presumably this can always be done, but in some cases it is more convenient to do than in others, as will appear.

The resonant element may be any of a wide variety of forms, mechanical or electrical, all characterized by the single property that, if deformed from a rest condition and released, the stored energy is transformed back and forth from potential to kinetic at a rate depending chiefly on the effective mass and the effective stiffness, or other like properties, a small proportion

* Encycl. Brit. 14th Ed. "A clock consists of a train of wheels, actuated by a spring or weight or other means, and provided with an oscillating governing device which so regulates the speed as to render it uniform."

of the energy being lost in internal friction at each oscillation. Some resonant elements which have been used in timekeepers are illustrated in Fig. 3.

The simplest appearing of all these is that of a mass, M , supported by a spring with stiffness, S . From the equation of motion

$$Sx = M \frac{d^2x}{dt^2}$$

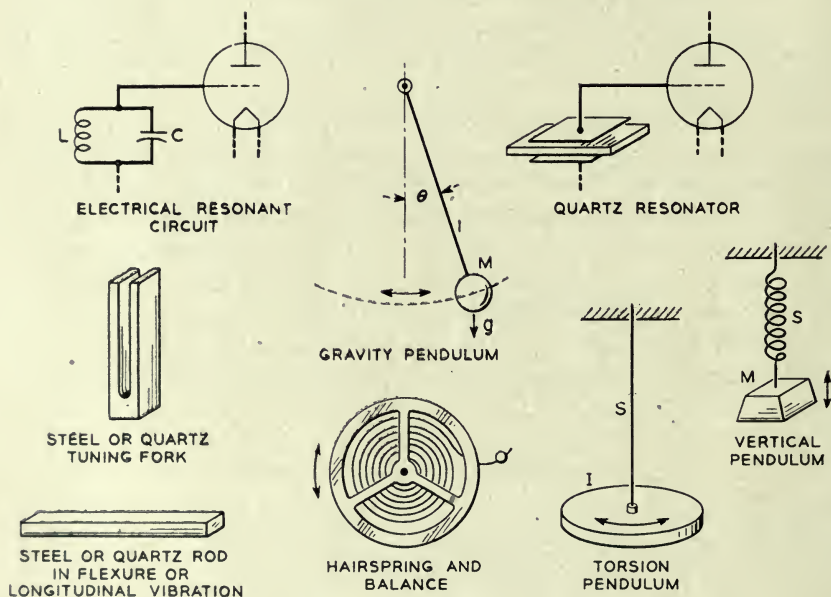


Fig. 3.—Typical resonant elements used in timekeeping.

the period of oscillation may be derived simply and is found to be

$$T = 2\pi \sqrt{\frac{M}{S}}$$

Similarly for the simple electrical resonant circuit where current flowing in an inductance, L , behaves like a mass, and current flowing in a condenser, C , behaves like the reciprocal of a stiffness, the period may be written.

$$T = 2\pi \sqrt{LC}$$

Similar expressions are derivable for the periods of oscillation of all simple oscillating systems, including the pendulum for which the period (for small amplitudes) is given by

$$T = 2\pi \sqrt{\frac{\ell}{g}}$$

where ℓ and g are respectively the length and gravity expressed in the same system of units, for example, the c.g.s. system.

When any such resonant element is strained from its rest condition, and released, it will oscillate with gradually decreasing amplitude until all of the stored energy has been dissipated in internal friction or resistance, and in the friction or resistance of the coupling with the supports. In general, the resulting amplitude of free oscillation may be given as

$$A = A_0 e^{-kt} \sin pt$$

the graph of which is a damped sine wave. The rate of free oscillation, p , is dependent chiefly on the effective mass and stiffness and to a small degree on the effective resistance of the element, while the rate of loss of amplitude, that is, the logarithmic decrement, k , is dependent on the ratio of effective resistance to effective mass.

If the resistance could be made exactly zero, such a motion once started would continue forever and its rate would be controlled wholly by the effective mass and stiffness of the resonant element. Actually, of course, such a condition cannot be realized in practice but, by the selection of suitable materials and environment, and by special control means, it is possible to approach very closely to the ideal condition by causing the oscillation to be maintained *almost* as though there were no damping.

The evolution of precision timekeeping, whether consciously or not, has centered around the study and development of these two ideas: to discover resonant elements whose rate-determining properties are inherently stable, and to discover means for sustaining them in oscillation as though they had no effective resistance; or in employing means to circumvent or to compensate for any such resistance. The high precision of rate control that can now be obtained has been the result largely of developments in these two categories.

The Pendulum

The gravity pendulum was the first truly resonant element to be used to regulate the rate of a clock and for nearly three centuries maintained the supremacy for precision measurements of time. The pendulum was more a discovery than an invention, the popular story of its origin being that, while still a youth of seventeen years, Gallileo Galilei chanced to notice that a hanging lamp in the Cathedral of Pisa seemed to swing at the same rate regardless of amplitude. This he confirmed approximately by comparison with his pulse, and later made an extensive study of the isochronism of swinging bodies. These studies were in progress as early as 1583. Nearly sixty years later Gallileo described to his son Vincenzio how a pendulum could be used to control a clock, but no concrete result of this advice is known to have been made at that time. A working model of this clock,

made subsequently from the original drawings, is on exhibition in the South Kensington Science Museum, London. The first authentic record of the actual use of a pendulum in a clock is attributed to the great Dutch scientist, Christian Huygens, who produced his first pendulum clock in 1657. This was described by him in the *Horologium* in 1658.²

The performance of pendulum clocks was so good that almost immediately clocks of all other types were modified to include a pendulum. So complete was this transformation that very few unmodified clocks are now in existence which antedate the first application of the pendulum to time-keeping. This, as a matter of fact, is one of the major reasons that so little is known about the actual mechanisms used in mechanical clocks that were made before the introduction of the pendulum.

The subsequent history of pendulum clock development is well described in numerous books and papers and covers a wide field. Only those factors that relate the pendulum to other means of rate control will be discussed in the following.

The properties of a pendulum which make it such a good timekeeper are easily seen from a study of the forces on the bob as illustrated in Fig. 3. Since these forces must be in equilibrium at all times we may write (assuming no friction)

$$Mg \sin \theta = M\ell \frac{d^2\theta}{dt^2}$$

The nearly isochronous property of the pendulum is contained in this relationship since the period, on solution, is

$$T = 2\pi \sqrt{\frac{\ell}{g}} \left(1 + \frac{1}{4} \sin^2 \frac{\theta}{2} + \frac{9}{64} \sin^4 \frac{\theta}{2} + \dots \right)$$

where θ is the maximum semi-amplitude of swing expressed in radians. When this arc is small the period approaches a minimum. For small angles the natural period depends almost wholly on the ratio of ℓ to g and the stability of T depends chiefly upon the constancy of ℓ and g . Figure 4 shows the relation between period and the arc of swing, expressed as seconds per day departure from the theoretical rate for zero arc.

The sum of all the terms that depend upon powers of $\sin \theta/2$ is known as the circular error, relating to the fact that the bob is constrained to move on the arc of a circle. It was shown theoretically by Christian Huygens³ that if the bob could be constrained to move on the arc of an epicycloid it would be truly isochronous, that is, the period would be completely independent of its amplitude of motion. It is of interest to note at this point that in no other resonator used for precision timekeeping is there

the direct counterpart of circular error, for in all other cases the restoring force varies linearly with displacement in the region of operation and not as a sine function of it.

In the early stages of pendulum clock development it was not necessary to consider the arc error because other errors were of greater magnitude. But it is by no means a negligible factor, and in all precision timing by pendulums it must be accounted for, either by allowing for an arc correction, as is done commonly in geodetic survey work, or by keeping the arc small and precisely controlling it. According to F. Hope-Jones⁴, referring to the master pendulum in the famous Synchronome free-pendulum clock: "A variation of only 0.01 mm. in the excursion of the bob or 2 secs. of arc will by circular error alter the rate by 0.00145 sec. per day,—and if it arose un-

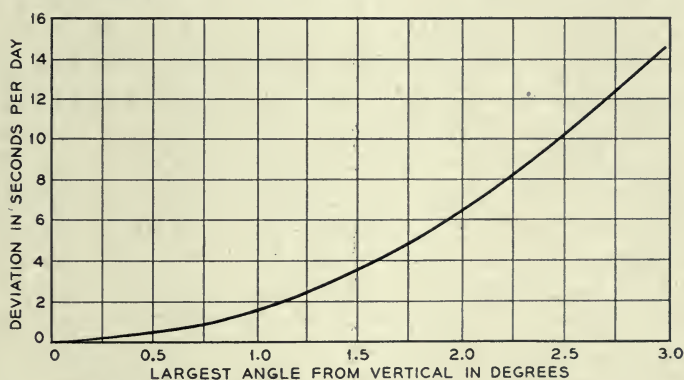


Fig. 4—Relation between arc and rate of pendulum.

perceived and was steadily maintained, it would produce an accumulated error of half a second in a year, so the necessity for this close observation is obvious."

The control of arc has almost invariably been accomplished by keeping constant the amount of energy applied per swing so that the actual amplitude obtained is that value for which all of the applied energy is dissipated in the pendulum system. In a sense this method of control of arc puts a penalty on improvements in design that would reduce the friction, because the better a pendulum becomes in this respect the less stable becomes the arc control. Since even the best pendulums develop unexplainable small changes in arc, it has been common practice in some observatories to record the arc frequently and to make allowance for changes in it when making the most precise time determinations.

The inherent constancy of rate of a pendulum, with small or constant amplitude of swing, depends to the one-half power on the stability of ℓ/g .

The changes in ℓ and g are quite independent of each other and so can be treated separately. Other factors that will be described also affect the rate, and it is the object in every precision clock design to reduce such variable effects to the absolute minimum.

Some control can be exercised over every factor except g , which remains a property of space and is dependent only on the proximity of matter and on acceleration. As is well known, the value of g varies over the surface of the earth due chiefly to its deviation from spherical shape, and because of the uneven distribution of matter. It also varies with vertical displacement or tides at any location to such an extent that a gravity clock that keeps accurate time at ground level will lose a second a day or more in a tall building. Actually, it is now possible to chart variations in g with high precision through measurement of the rate of a pendulum clock against a standard whose rate does not depend upon gravity.

Most of the factors that can affect ℓ have been studied critically and means have been found to reduce them to very small effects. The chief source of variation was at first the temperature coefficient of the pendulum rod. With ordinary metals the rod expands from 10 to 16 parts in a million per degree C, causing a proportionate change in rate of half this amount, corresponding to from one-half to two-thirds of a second per day. Many ingenious means were developed to reduce this effect, starting with George Graham's mercury-filled bob in 1721, followed by John Harrison's grid-iron pendulum in 1726, and a great number of variations on these ideas, all depending on the differential coefficient of expansion of dissimilar materials.

About the year 1895, Charles Edouard Guillaume of Paris developed an alloy, consisting chiefly of nickel and iron, which he called Invar, because it had a very small temperature coefficient of expansion, from which pendulum rods could be made. The use of this material made it unnecessary to resort to complex compensated pendulums with their own inherent instabilities, and the accuracy of timekeeping was increased another step. The residual temperature effects could be measured readily, and compensated if desired, by the use of a small bar of aluminum attached to the bob.

Some other important factors that affect the working length of a pendulum are the aging of the supporting rod, the "knife edge" or spring used for the suspension, the nature of the main supporting column or frame, and some atmospheric effects caused by changing temperature and pressure. In the most accurate pendulum clocks, the atmospheric effects are greatly reduced by mounting the pendulum in partially evacuated, hermetically sealed enclosures which can be temperature controlled. All of these factors and many others are discussed in every good treatise on accurate pendulum clocks. They are mentioned here chiefly for the purpose of comparison with like factors in the quartz crystal clock and to show how in many

cases the difficulties introduced by such factors may be more easily and more positively controlled.

In every primary clock mechanism the resonant governing device must be sustained in oscillation, and the manner in which this is done has a strong bearing on its rate regardless of the quality of the governing element. The basic requirements are the same for any kind of oscillator, whether a pendulum, an electrically resonant circuit comprising inductance and capacitance, a steel tuning fork, or a quartz crystal resonator. The requirements were first stated for the case of the pendulum by Sir George Airy in 1827 and it has always been the aim in the design of every good pendulum driving means to satisfy Airy's condition.

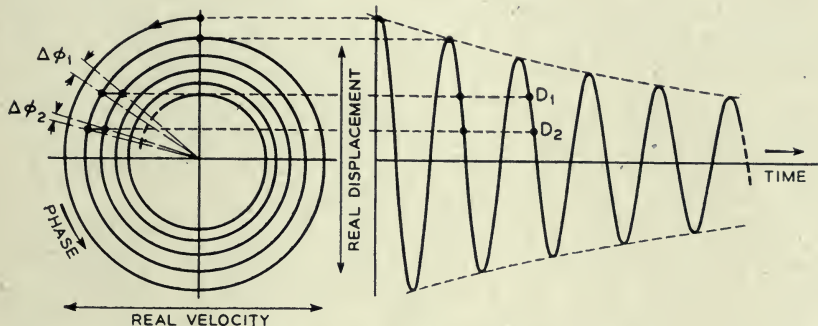


Fig. 5—Amplitude-phase diagram for resonant element.

This condition is conveniently illustrated by the diagram of Fig. 5 which shows the two most familiar representations of damped sinusoidal motion. In order to provide a convenient scale in the drawing an impractically large damping is represented, corresponding to a Q of 20. The Q of a resonant circuit is related to the logarithmic decrement, δ , by the relation $Q\delta = \pi$. The factor δ is the logarithm, to base $e = 2.718 \dots$, of the ratio of the amplitudes at any two successive periods. It should be noted that the Q of a good electrically resonant circuit is in the order of 200, that of a good pendulum from 10,000 to 100,000 and that of a good quartz resonator from 100,000 to 5,000,000. The significance of these higher values of Q will be evident from the following discussion.

In Fig. 5 the damped sine wave shown corresponds, point by point, to the phase diagram, which is simply a logarithmic spiral. By suitable choice of scale the spiral can be interpreted to represent either the amplitude or the velocity—in which case the real amplitude is vertical and the real velocity horizontal. In this representation the velocity is shown maximum when the amplitude is zero, which is a very close approximation to fact for all practicable values of Q . The discussion will center on the velocity spiral.

Let us assume that the pendulum is sustained in oscillation by a succession of short impulses, one for each swing applied at some phase angle ϕ_1 . If the impulse is really short, the *velocity* will be increased to the value that the pendulum had when it occupied the same *position* during the last swing. This change of condition is represented by the short horizontal path on the velocity-phase diagram and, as indicated, is accompanied by an advance in phase $\Delta\phi_1$. This can be interpreted as meaning that the period of a pendulum sustained in oscillation in this way is reduced from its natural period in the ratio of $\frac{2\pi - \Delta\phi_1}{2\pi}$. It is obvious from the diagram that $\Delta\phi_1$ becomes

smaller and that this ratio approaches unity as the phase of the applied impulse approaches that of the maximum velocity—that is, when the pendulum is in the center of its swing; and this is Airy's condition. It is clear also that if the impulse is applied after (instead of before) the instant of maximum velocity, the period will be correspondingly increased. From the geometry of the figure, it can be seen that, in the neighborhood of the optimum condition; the deviation from natural period is very closely proportional to the amount of the phase departure.

The closeness of spacing of the turns of the spiral depends directly on the Q of the resonant element. For a Q of 200, the turns will be packed ten times closer than shown, and the corresponding $\Delta\phi$ will be only one tenth as great, other conditions being comparable. For a Q of a million or more, $\Delta\phi$ becomes very small indeed, especially when ϕ is properly chosen—and the *variation* in $\Delta\phi$, which is a measure of the variation in rate due to the driving means, may be made vanishingly small.

The importance of the above properties to timekeeping depends upon how well conditions can be set up to realize them. At first wholly mechanical means were employed and, with the advent of the dead-beat and detached escapements and by careful design and operation, quite remarkable performance was obtained.

A new approach in timekeeping methods was introduced by Alexander Bain⁵ in 1840 when he first used electrical means for sustaining a pendulum in oscillation. The importance of Bain's invention of the electric clock is indicated by a long controversy over the priority of the invention with Charles Wheatstone, who was working along similar lines at the same time as a by-product of his extensive researches on the electrical telegraph. A brief story of this controversy entitled "The First Electric Clock" was written for the one-hundredth anniversary of Bain's invention⁶. The first electric pendulum clocks could not compare in accuracy with the best mechanically driven pendulums of the period but, in spite of a great deal of initial skepticism on the part of those brought up in the mechanical

tradition, electrical maintenance and control has been applied in the most accurate pendulums in the world.

The free-pendulum clock makes use of the idea, first proposed by Rudd, of allowing a master pendulum to swing free of all sustaining or other mechanism for a considerable number of periods and of imparting to it, after each group of free swings, a single impulse large enough to maintain the next equal number. The advantage is that no friction effects of driving mechanism are coupled to the pendulum except during that minimum time required to impart energy to it. Actually, in theory, the phase error introduced by one large impulse after n free swings is exactly the same as the sum of the phase errors for n small impulses. That can be deduced from the phase diagram of Fig. 5. But experience has shown that a pendulum is actually more stable when the sustaining mechanism is detached from it the greater part of the time.

The Synchronome free-pendulum clock includes also the basic idea of the gravity remontoir first applied by Lord Grimthorpe (then Sir Edmund Beckett Denison) in the design of the mechanism of Big Ben, London, constructed in 1854—and still in continuous operation. The ingenious application of these principles and the electrical means devised by F. Hope-Jones and W. H. Shortt for its accomplishment have resulted in the construction of the most accurate pendulum clocks in the world by the Synchronome Clock Company of London. The history and development of the free-pendulum clock is elegantly described by F. Hope-Jones in his book on *Electric Clocks*⁷.

The predominant characteristics of a pendulum resonator, as used in a clock, have just been discussed in order to show the parallel between them and the properties of other resonant systems. It will be shown how some of the factors that have been troublesome in the development of pendulums have been rather easily taken account of in other types of control devices and in particular in the quartz crystal clock.

THE EVOLUTION OF ELECTRIC OSCILLATOR CLOCKS

It almost never happens that a result of any considerable value is obtained at a single stroke or comes through the efforts of a single person. More often even the most important advances come as the climax of a long series of ideas which have accumulated over a period of years until the next step becomes almost self-evident and is accomplished either through the necessity for a new result or as a logical next step.

This was preeminently the case in the crystal clock development and involved the putting together of a considerable number of ideas that had been accumulating through a century or more of related activity. The

chain of events which led eventually to the crystal clock followed a course quite independent of pendulum clock development, although parallel with it, and meeting it from time to time on the way. From the start, it involved the use of resonant elements whose frequencies do not depend upon gravity for controlling the frequency of oscillations in a positive feedback amplifier. From a rather simple beginning, taking advantage of a series of discoveries and inventions through about a century of progress, there has evolved a clock whose stability is comparable with that of astronomical time itself, as heretofore defined in terms of the earth's rotation, and having a versatility far exceeding all other existing means for the precision measurement of time.

Electric Oscillators

The first recorded experiments that relate directly to this development were those of Jules Lissajous⁸ who, in 1857, showed that a tuning fork can be sustained in vibration indefinitely by electrical means, using an electromagnet and an interrupter supported by one of the prongs. The idea of using an interrupter to sustain vibration was not new with Lissajous, but had been invented by C. G. Page⁹ and described by him as early as April 1837, to obtain a regularly interrupted electric current. Credit for this important invention is often given to Golding Bird¹⁰ or Neeff¹¹ who evidently were working along similar lines concurrently although quite independently of each other. Page, Golding Bird and Neeff were all medical doctors and evidently were interested in their devices more for their therapeutic interest than for the general scientific value, since "galvanic" electricity was attributed at that time with marvelous healing powers.

Lissajous was probably the first to make use of the idea for accurate measurements of rate, being a prolific experimenter in mechanics and acoustics, and the originator of the famous method bearing his name for the study of periodic motions. Indeed, the electrically operated fork was developed especially for use as a standard to be used in studying the rates of other vibrators. In principle, the electrically operated fork is like the pendulum drive of Alexander Bain, except that the rate of vibration in this case is not a function of gravity but for the most part is controlled by the effective mass and elastic stiffness of the vibrating member.

The tuning fork itself was invented in 1711 by John Shore, a trumpeter in Handel's orchestra¹², and was developed to a high state of perfection by the great instrument maker and physicist of Paris, Rudolph König. To establish an accurate standard of pitch for calibrating these forks König developed what he termed an "absolute" method for the determination of frequency. This consisted of a tuning fork having a frequency of 64 vibra-

tions per second, with delicate mechanical means, similar to a clock escapement, for sustaining the fork in vibration and for counting the number of vibrations over any desired interval of time. For this purpose, the escapement mechanism was geared to the hands of a clock, so that when the fork had its nominal frequency the clock would keep correct time. Dr. König credits the invention of the fork-clock to N. Niaudet¹³ in these words:

“Cette disposition avait été réalisée pour la première fois dans l’horloge à diapason que N. Niaudet fit présenter à l’Académie des Sciences le 10 décembre 1866, et que à figuré aux expositions universelles de Paris 1867 et de Vienne 1873.”*

Thus, as early as 1866, the essential elements had been developed separately from which a clock of the electric oscillator type could have been constructed. But it was not until more than half a century later, when there was more apparent need for such a clock, that it was actually realized. It was chiefly for the purpose of studying temperature coefficients and like properties of tuning forks that König constructed and used his famous mechanical fork-clock. There is no evidence that there was at that time any idea of using a fork-clock as a timekeeper.

It was for the purpose of making still more precise studies of the properties of tuning forks that H. M. Dadourian¹⁴ in 1919 made use of the phonic wheel motor for the first time for counting the number of cycles executed by a fork over an extended period of time to measure its rate. By means of a chronograph the time interval corresponding to the total of a very large number of periods could be measured precisely in terms of a standard clock, thus providing a direct “absolute” measure of fork rate. For this he found already invented for him all of the essential component parts, including the fork with electromagnetic drive, and the phonic wheel motor.

The phonic wheel motor, which in some modified form is an essential part of nearly all oscillator clocks, was invented by two investigators, apparently quite independently and for entirely different purposes. The first published reports of each appeared in 1878.

The first of these is an American patent that was granted on May 7, 1878 to Poul La Cour¹⁵, a Danish telegraph engineer. The application was filed in Washington on April 9 of the same year, and described a fork-controlled impulse motor similar to those still used in many modern synchronous clocks. The other publication was a report in *Nature* for May 23 of the March 30 Physical Society Meeting. In this, Lord Rayleigh described a motor which he developed to measure the frequency of sound by a stroboscopic method.¹⁶ Both of these original disclosures indicated a

* “This apparatus was realized for the first time in the fork-clock which N. Niaudet described at the Academy of Sciences on December 10, 1866, and which was shown at the expositions of the University of Paris in 1867 and the University of Vienna in 1873.”

considerable amount of previous study, even including the fluid-filled flywheel to reduce hunting. It may be impossible at this time to know who actually put in motion the first phonic wheel motor.

Difficulties inherent to contact-controlled devices prevented the development of highly accurate fork standards of this type, and there is no evidence so far that any thought had been given to the use of a tuning fork as a timekeeper.

The method of using a microphone instead of a contact was proposed by A. and V. Guillet¹⁷, in 1900 and has been used considerably in frequency standards of moderate accuracy, but that too had limitations which made it impossible to utilize fully the inherent stability of a good tuning fork.

The Use of Vacuum Tubes

The first opportunity for really precise control of the frequency of a mechanical vibrating system, and the next step in the oscillator clock evolution, came with the invention of the thermionic vacuum tube at the turn of the century. The development of the vacuum tube has been a more or less continuous process¹⁸ starting with the studies of electrical conduction in the neighborhood of hot bodies by Elster and Geitel, Edison, and Fleming, and later developed into the first practical devices by Fleming¹⁹ and DeForest²⁰ in England and America respectively. The first patent for such a device, a two-element tube, was issued to J. A. Fleming in 1904.²¹ The first patent on a tube containing three elements and suitable for use as an amplifier was issued to Lee DeForest in 1907.²²

The vacuum tube as an amplifier found almost immediate and widespread application in telephony and, next to the basic telephone elements, was the most important single factor contributing to long distance communication. For this purpose large amounts of amplification were required. Very often in the operation of early amplifiers, enough signal from the output would somehow get coupled into the input circuit to make the entire circuit break into oscillation on its own account at some frequency for which the amplifier and feedback circuit were particularly efficient.

Although this was very annoying in an amplifier, it led naturally in 1912 to the invention of the vacuum tube oscillator, consisting essentially of an amplifier with coupling between the output and the input and some definite means for regulating the frequency of oscillation. The first to seek patent protection in vacuum tube oscillators were Siegmund Strauss²³ in Austria, Marconi Company in England²⁴, A. Meissner in Germany, and Irving Langmuir, E. H. Armstrong and Lee DeForest²⁵ in America. Many specific forms have since been invented and widely used, some of the more familiar types being associated with the names of Colpitts, Hartley and Meissner.

With the vacuum tube oscillator controlled by electric circuit elements, it would have been possible immediately to operate a clock by means of a phonic wheel motor. Even if this had been done, however, the accuracy would not have compared very favorably with that of good mechanical clocks of the period. This is because the rate-controlling element of such oscillators was subject to large changes due to temperature and aging, and because means were not yet known for avoiding the effects of tube and other variables on the resulting frequency.

The next important step in our evolution was the use of the vacuum tube to sustain the vibration of a tuning fork. This may be considered either as an improvement on the contact-driven fork by the substitution of a vacuum tube relay device instead of the contact, or as an improvement on the vacuum tube oscillator by the substitution of a mechanical resonator for the electrical resonant element. This achievement was first announced by Professor W. H. Eccles²⁶ in April or May, 1919, and was followed on June 20 by a note by Eccles and Jordan²⁷ in the *London Electrician*. Meanwhile, on June 16 of the same year, a similar announcement appeared in *Comptes Rendus* by Henri Abraham and Eugene Block²⁸, showing that parallel developments were in progress in both England and France. However, Eccles and Jordan in discussing their work at the National Physical Laboratory stated: "Several instruments of this kind have been set up and used during the past 18 months." From this, we may imply that they had vacuum tube driven forks in operation early in 1918.

One of the chief advantages of the use of the vacuum tube to sustain oscillations in a mechanical system is that the variable friction of the contact mechanism is avoided. Previously this had been one of the main causes of instability. With the new method it became possible to operate in a wide frequency range, continuously, and at small amplitude, and to deliver alternating currents of approximately sine wave form and having more constant frequency than heretofore had been possible. The judicious use of a vacuum tube in delivering power to sustain the vibration of a resonator is analogous to the ideal of the so-called free pendulum but may be utilized more effectively in freeing the resonator from disturbing influences associated with the driving means.

Another important advantage, which, however, was not realized immediately, is the ease with which the phase of the driving force applied to a mechanical vibrator can be adjusted for greatest frequency stability. In a manner analogous to the pendulum, in which it was shown that the rate is least affected when the driving impulse is applied at the instant of maximum velocity, the current delivered to the driving electromagnet and hence the force applied to the vibrating element, should be in phase with

the *velocity* of that element. In the vacuum tube oscillator, it is a relatively simple matter to design the feedback circuits to meet this condition very accurately.

In 1921 and 1922 Eckhardt, Karcher and Keiser^{29, 30} described the development of a precise fork and vacuum tube driving means, pointing out the following uses: "As a sound source; as a small scale time standard; as a current interrupter; as a synchronizer." The chief emphasis seems to have been on the second item because in the same year Eckhardt described a high-speed oscillograph camera using the same fork as a precise timing device. The study and improvement of the tuning fork oscillator were carried on continuously and soon such oscillators were used in several national physical laboratories and commercial research institutions as standards of frequency and time interval.

The next two reports of progress appeared in 1923, one by D. W. Dye of the National Physical Laboratory in Teddington, and the other by J. W. Horton, N. H. Ricker and W. A. Marrison of Bell Telephone Laboratories, New York City. Both of these papers disclosed work done over a period of two or three years and described apparatus that had been in operation for a considerable period. Dr. Dye employed a 1000-cycle steel tuning fork and a phonic wheel motor operating synchronously from it with a gear reduction and cam to produce periodic electrical signals which he compared with a clock by means of a chronograph³¹. Horton, Ricker, and Marrison used a 100-cycle steel fork, a synchronous motor with a gear reduction to produce electrical impulses at one-second intervals, and a clock mechanism operating directly from these signals³². This appears to be the first time that a vacuum tube-controlled oscillator was ever used to operate a complete clock mechanism. Shortly thereafter, a clock was built in which the 100-cycle motor was geared directly to the clock mechanism instead of operating through a stepping device. A contacting device was retained, however, for the purpose of making precise time measurements.

For precise measurements of rate over long time intervals, means were provided to compare the seconds pulses controlled by the synchronous motor directly with time signals received by radio from the Naval Observatory. To facilitate these comparisons, a two-pen siphon recorder was built by means of which the time marks were laid down side by side on a moving strip of paper in such a way that accurate subdivisions of a second could be made on any part of the record.

This same two-pen recorder and 100-cycle fork time standard was used during the total solar eclipse of January 24, 1925 to time the progress of the moon's shadow as observed at a number of stations in the path which were all connected by a round-robin telegraph circuit, through the Bell

Telephone Laboratories' headquarters in New York City^{33,34}. A photograph of the original records is reproduced in Fig. 6. This is believed to be the first time that a vacuum tube oscillator type of time standard was ever used in the service of astronomy.

During the following ten years a great number of improvements were made in tuning fork oscillators and they became widely used as precise frequency standards. The Bell Laboratories' 100-cycle fork standard was mounted in a container which could be sealed at constant pressure or vacuum. It was carefully temperature controlled and provision was made to keep the amplitude within prescribed limits. In describing this improved standard³⁵, comprising a synchronous motor geared directly to a clock mechanism, the authors Horton and Marrison made the following statement:

"During tests on this frequency standard, it was found that it constituted a far more reliable timekeeper than the electrically maintained pendulum clock which was used to obtain the data already published. The pendulum clock was, therefore, dispensed with and all measurements of the rate of the fork are now made by direct comparison with the mean solar day as defined by the radio time signals sent out by the U. S. Naval Observatory."

In all fairness to the pendulum clock in question, it should be stated that the laboratory was situated on the seventh floor of a building adjoining a busy street and so was continually subject to vibration from traffic, wind, and other changing conditions. Disturbances of this sort have little or no effect on standards of the electric oscillator type but seriously impair the performance of most high precision pendulum clocks. The relative immunity of the oscillator standard to change of position and shock has an important bearing on its value in many applications.

Probably the most precise tuning fork controlled time and frequency standards ever constructed were those developed in the National Physical Laboratory at Teddington, as a continuation of the work begun there by Professor Eccles and carried forward by Dr. Dye and his staff. A report by D. W. Dye and L. Essen in the Royal Society Proceedings in 1934³⁶ described a number of refinements in the fork and method of use some of which had been suggested by Dr. Dye as a result of his studies ten years earlier. Among these was the use of *elinvar* in the construction of the forks in order to reduce the effect of variable temperature on the frequency. Elinvar is a nickel steel containing about twelve per cent of chromium, which on proper treatment has a small or zero temperature coefficient of elasticity. It was invented by Charles Edouard Guillaume^{37,38} and was further studied by P. Chévenard^{39, 40}. The excellence of the N.P.L. fork standard can be appreciated readily from the conclusion of the 1934 report which states in part:

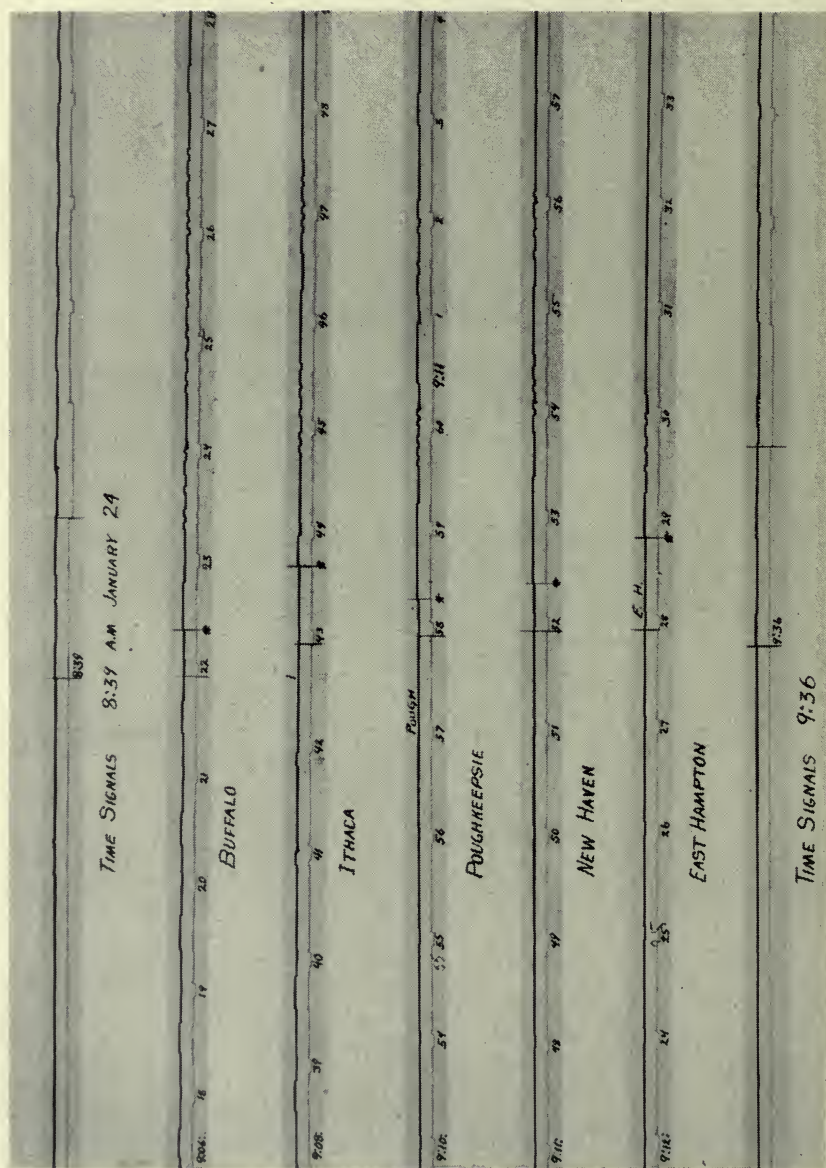


Fig. 6—Timing records of total solar eclipse of January 24, 1925.

"The frequency of the fork in comparison with the N.P.L. Shortt clock can be measured at any time with an accuracy of 5 parts in 10^8 . It is necessary to apply a correction for the rate of the Shortt clock, and the ultimate accuracy with which the absolute value of frequency is known depends on the accuracy of the time signals which are used to determine the rate of the clock. The final frequency can, however, usually be ascertained with an accuracy of ± 1.5 parts in 10^7 . In its present condition the tuning fork maintains a frequency stability of the order of 3 parts in 10^7 over periods of a week or more."

A considerable amount of effort has been devoted to the improvement of tuning forks, directed mostly toward stabilizing the fork itself. Patents issued to H. H. Hagland⁴¹, August Karolus⁴² and Bert Eisenhour⁴³ have been concerned with the reduction of temperature coefficient by various methods of compensation in the alloy or in the mechanical structure of the fork. In recent years, alloys have been produced from which forks with a zero coefficient of frequency can be machined. These alloys have neither a zero expansion coefficient nor a zero elastic coefficient, but the two coefficients are so balanced that their effects cancel as they concern the frequency of a tuning fork.

One of the largest residual sources of error in a good fork is that caused by the coupling through the mounting. A fork which is efficient as a producer of sound by coupling through the base would be quite useless as a precise standard of rate due to the losses introduced in this manner. It has been shown by S. E. Michaels⁴⁴ that the tines of a well-balanced fork can be so shaped that practically no energy at fundamental frequency is transmitted through the base.

By making use of all that is known about materials, shapes and mountings for tuning forks, and all that is known about stabilized vacuum tube circuits for driving them, it is quite possible that considerable further improvement could now be obtained in such a standard. But another line of development has shown greater promise in this field and the ultimate accuracy of tuning fork oscillators has never been pursued.

The Quartz Resonator

During the same ten years that the greatest advances were being made in the tuning fork art, the striking properties of the quartz crystal resonator were reviewed and first applied in the construction of frequency and time standards. Its use in primary standards for the most exacting measurements of frequency and time is now almost universal in national and industrial laboratories throughout the world.

Quartz crystal is the most abundant crystalline form of silicon dioxide, occurring, in some parts of the world, in large single crystals from which mechanical resonators of useful dimensions can readily be formed. The physical properties that make it eminently suitable for use in a standard of rate or time are its great mechanical and chemical stability. Having a

hardness nearly equal to that of ruby and sapphire, and a rigidity of structure such that it cannot be deformed beyond its elastic limit without fracture, it might be expected to remain in a given shape indefinitely under ordinary conditions of use. Because of its great chemical stability, its composition is not easily modified by any ordinary environment.

In addition to its inherent physical and chemical stability, the elastic hysteresis in quartz is extremely small. For this reason, it requires only a very small amount of energy to sustain oscillation and the period is only very slightly affected by variable external conditions in the means for driving it.

A striking illustration of the importance of this property is indicated by the number of periods that a resonant element will execute freely, that is, without any sustaining forces whatever, during the time required for the amplitude to decrease to one-half of some prescribed value. For a good electrical circuit consisting of an air core inductance and an air condenser, this number is about 100; for a good tuning fork in vacuum, it is about 2000. For a good cavity resonator under standard conditions of temperature and pressure, the number may be as high as 10,000. The best gravity pendulums will swing freely from 2,000 to 20,000 times before they reach half amplitude. The effect is most striking of all in quartz crystal, in which the internal losses are extremely low. Professor Van Dyke has measured the rate of decay of oscillations under a wide range of conditions⁴⁵ and has found that, as ordinarily mounted, nearly all of the losses are in the mounting or in the surrounding atmosphere, if any, or in surface effects. Extremely small amounts of surface contamination will more than double the decrement. Recently⁴⁶ Maynard Waltz and K. S. Van Dyke have measured the decrement of one out of the first set of four zero coefficient ring crystals ever made⁴⁷ and found that, vibrating freely in vacuum and favorably mounted, it would execute more than a million vibrations before falling to half amplitude.

The advantage of this property is immediately obvious because of the relatively small amount of energy that must be supplied at each oscillation to keep the resonator in motion. As already discussed in relation to the pendulum, the amount that the rate of oscillation may be disturbed in a given structure is proportional to this energy and, to first order, on the departure from the ideal phase condition of the applied driving force.

The properties just enumerated are sufficient to assure the superiority of quartz crystal for the control element in a rate standard; no other vibrating system known at the present time is so sharply resonant or so stable. However, one more property, its piezoelectric activity, has added greatly to the convenience of its use in vacuum tube devices.

The piezoelectric effect was discovered by the Curie brothers in 1880,⁴⁸

and in the years following was studied extensively by them^{49, 50}. They found that when quartz and certain other crystals are stressed, an electric potential is induced in nearby conductors and, conversely, that when such crystals are placed in an electric field, they are deformed a small amount proportional to the strength and polarity of that field. The first of these effects is known as the *direct* piezoelectric effect and the latter as the *inverse* effect. The amount of such deformation in quartz is extremely minute, a static potential gradient of 1 esu (300 volts) per centimeter causing a maximum extension or contraction, depending on the polarity, of only 6.8×10^{-8} cm per cm. If a crystal resonator is subjected to an alternating electric field having the frequency for which the crystal is resonant, the amplitude of motion will, of course, be multiplied many times. In practice, however, the actual amplitudes of motion are kept so small, by limiting the applied electric field, that even with the largest crystals used they can

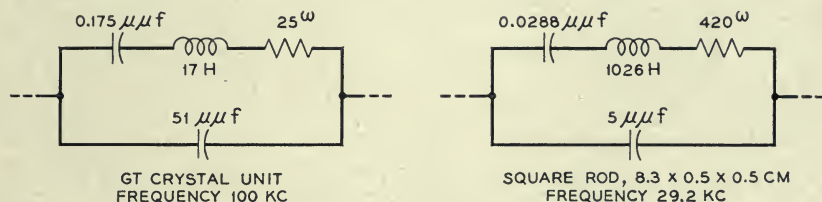


Fig. 7—Equivalent electrical circuits for typical quartz crystal resonators.

be observed only under a high powered microscope. This, in conjunction with means for precise amplitude control, is one of the reasons for the remarkable frequency stability of quartz crystal oscillators.

In practice, a quartz resonator is mounted between conducting electrodes which now most often consist of thin metallic coatings deposited on the surface of the crystal by evaporation, chemical deposition or other suitable means. Electrical connection is made to these coatings through leads which also support the crystal mechanically. The resonators with which we are chiefly concerned in this discussion have only two electrodes.

If such a two-terminal resonator is connected into any circuit, it will behave there *as though* it consisted of wholly electrical circuit elements, usually of such low loss as can not be realized by other means. The equivalent electric circuit for a quartz crystal resonator was first described⁵¹ by K. S. Van Dyke in 1925 and, for some significant cases, is illustrated in Fig. 7. The part of such an equivalent circuit which in many cases cannot be duplicated by any ordinary means is the inductance element containing so little resistance. It is as though an electric resonator could be made and utilized constructed of some supra-conducting material.

Among the first serious efforts to utilize the piezoelectric effect in electrical circuits were those of Alexander McLean Nicolson who used rochelle salt crystal in the construction of devices for the conversion of electrical energy into sound and vice versa. He constructed loudspeakers and microphones during several years of study prior to the publication of his work⁵² in 1919—ideas now being used extensively in phonograph pickups, microphones and sound producers. Nicolson also was the first to use a piezo-active crystal to control the frequency of an oscillator. His patent⁵³, applied for in 1918, shows a circuit which he operated successfully in 1917. The first actual use of resonators of quartz is attributed to P. Langevin^{54, 55}, who drove large crystals in resonance in order to generate high-frequency sound waves in water for submarine signaling and depth sounding.

The Quartz Crystal Controlled Oscillator

The first comprehensive study of the use of quartz crystal resonators to control the frequency of vacuum tube oscillators was made by Walter G. Cady in 1921 and published by him in April, 1922⁵⁶. This was the step which initiated a most extensive and intensive research of the properties of quartz crystal and into methods for its use in numerous fields requiring a stable frequency characteristic.

The extent and importance of this research are well indicated by the number of investigators and published contributions to the art. Among these, a paper by A. Scheibe⁵⁷ in 1926 lists 28 articles on the subject, along with a description of his own extensive studies. Two years later Cady published a bibliography⁵⁸ on the subject, including 229 separate references to papers and books and 84 patents in various countries. R. Bechmann in 1936 published a review of the quartz oscillator⁵⁹ including 26 references to other original contributions in that field alone. More recently there comes at the end of Cady's 1946 book⁶⁰ on "Piezoelectricity", a bibliography of 57 books and 602 separate published articles on this subject. By any measure this represents a great amount of detailed effort for a single subject in so short a time—just about a quarter of a century. Of this great amount of material, it is feasible to review only a small number of the outstanding ideas relative to the evolution of the quartz crystal clock.

The first published quartz-controlled oscillator circuit is reproduced in Fig. 8A from Cady's 1922 article. In this oscillator the "direct" and "inverse" piezoelectric effects were employed separately, making use of two separate pairs of electrodes. The output of a three-stage amplifier was used to drive a rod-shaped crystal at its natural frequency through one pair of electrodes making use of the "inverse" effect, while the input to the amplifier was provided through the "direct" effect from the other pair.

The feedback to sustain oscillations in the electrical circuit could be obtained only through the vibration of the quartz rod and hence was precisely controlled by it. Cady's results were received with widespread interest and were duplicated and continued in many laboratories, which soon resulted in many new discoveries and inventions.

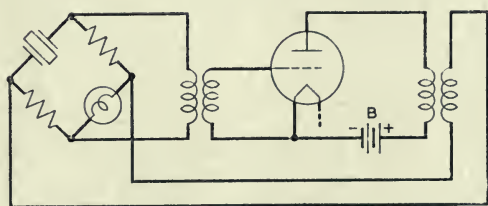
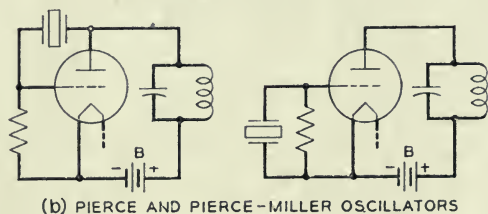
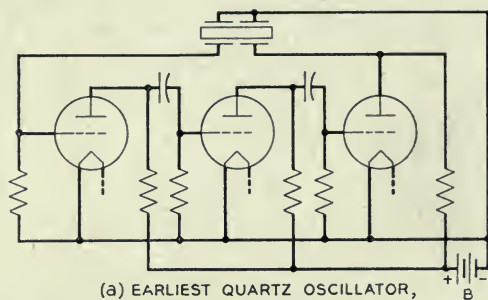


Fig. 8—Typical quartz oscillator circuits.

Important contributions were made by G. W. Pierce, who, showed in the following year that plates of quartz cut in a certain way could be made to vibrate so as to control frequencies proportional to their thickness⁶¹. He also proposed somewhat simplified circuits for their use which soon found very general application in the construction of wavemeter standards and later for oscillators used to control the frequency of broadcasting stations and for many other purposes. In 1924, the General Radio Company of Cambridge, Massachusetts, produced a commercial instrument based on these studies.

The significance of the unusually stable properties of quartz crystal—which at times were viewed with a sort of awe and a tendency at first to expect too much^{62*}—was soon recognized in relation to precise standards of frequency and time, and many laboratories made experiments directed toward these applications.

For some years these efforts usually took one of two forms: either that of a quartz-controlled oscillator used as a comparison standard by various means⁶³, or that of using the quartz resonator itself as a portable standard, the high-frequency counterpart of an isolated tuning fork. Probably the most convenient standards of the latter sort were the luminous resonators first described in 1925 by Giebe and Scheibe⁶⁴. The following year they proposed the use of such luminous resonators as frequency standards⁶⁵ and, shortly following, portable frequency indicators of this sort were made available for general use. The use of such a luminous resonator for the international comparison of frequency standards was reported by S. Jimbo in 1930.⁶⁶ The first international comparison of frequency standards making use of piezo resonators as isolated standards was carried out by Walter G. Cady in 1923, who by means of a set of early type resonators compared the existing standards at Rome, Livorno, Paris, Teddington, Farnborough, Washington, and Cruft Laboratory at Harvard University⁶⁷. In the following year the U. S. Bureau of Standards carried out a similar international frequency comparison, but of greater accuracy, employing portable quartz crystal oscillators. This comparison and other important related studies were described by J. H. Dellinger in 1928—"The Status of Frequency Standardization"⁶⁸.

It was soon recognized that quartz oscillators could be built with a stability far greater than that of any other known type and that they possess qualities very desirable for a combined time and frequency standard. However, all early quartz oscillators had frequencies far too high to operate any synchronous motor and it was not immediately obvious how a clock could be operated thereby.

The Frequency Divider

The illustration in Fig. 9 from the author's notebook for November, 1924 is believed to be among the earliest means proposed to accomplish this. In brief, the proposal was to control the speed of a motor driving a high-frequency generator so that a harmonic of the generator output, say the

* In 1929, M. G. Siadbei wrote "Nous pensons que le quartz piécoélectrique peut trouver un nouvel emploi dans la chronometrie, étant donnée la conservation rigoureusement constant de ses oscillations."

"La seule cause de variation de la période d'oscillation résulte en effet du changement de la température...."

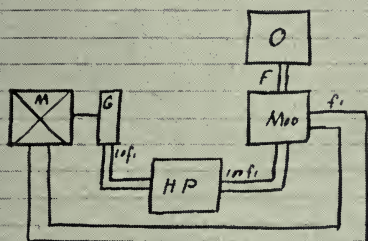
tenth, would have a frequency of the same order as that of the crystal but differing from it by a relatively low frequency, f_1 . This low frequency, derived from the modulator was to be used to drive the synchronous motor.

DATE November 15 1924.

55

Means for synchronizing a rotating machine with a current at radio frequency.

see p 177
also p 168 No 2161



Let M be a motor-generator
a generator G which supplies
current at 10 times the motor
input frequency f_1 . H.P. is
a harmonic producer giving
say the 10th harmonic. O is
an oscillator supplying current
at frequency F. The modulator
produces one component of
frequency $F - 10f_1 = f_1$.

Fig 1. Now let f_1 drive the synchronous motor M and we have a device for maintaining a shaft speed some rational multiple of the frequency of O.

The oscillator O may be a quartz crystal controlled oscillator having a frequency of 100,000 v or higher so that a very convenient method is provided for maintaining a radio frequency standard.

Motor M could be geared to a clock or any suitable recording device to facilitate checking frequencies.

W. D. Morrison November 15 1924.
S. G. Schellum November 15, 1924

Fig. 9—Early suggestion of means to control a rotating device such as a clock from a high frequency.

The shaft speed of the motor-generator would, therefore, be integrally related to the crystal frequency and hence any mechanism geared to the shaft, such as a clock, would indicate time as dictated by the crystal. This method could have been carried through readily by a combination of means already developed for other purposes, and the construction of an apparatus based on this suggestion was soon begun. However, a simpler method⁶⁹,

not involving a rotating machine in the control system, was suggested and the first quartz crystal clock was constructed using the simpler means. This apparatus was described by Horton and Marrison⁷⁰ before the International Union of Scientific Radio Telegraphy in October, 1927. The reso-

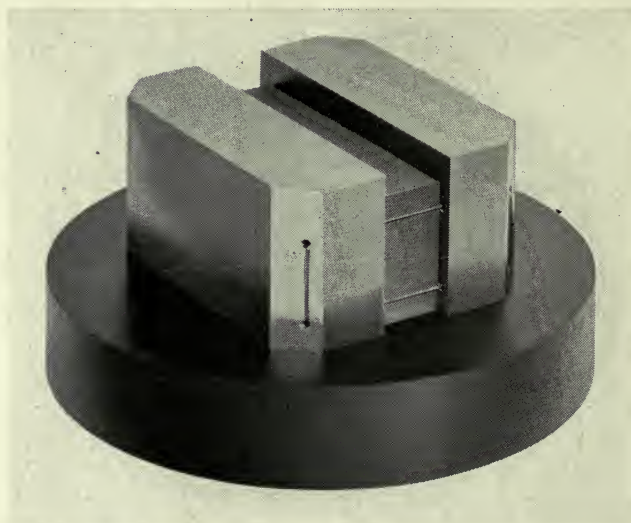


Fig. 10—50,000-Cycle quartz resonator, in original mounting, used in first quartz clock—1927.

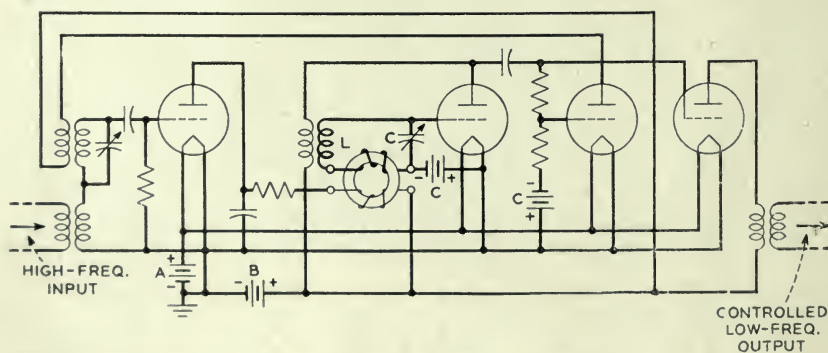


Fig. 11—Submultiple controlled frequency generator used in first quartz clock.

nator in its mounting that was used in this first model is shown in Fig. 10. It consisted of a rectangular block of crystal, cut in the manner usually called X-cut, and of such size as to oscillate at a frequency of 50,000 cycles per second in the direction of its length. The temperature coefficient of this resonator was approximately 4 parts in a million per degree C at the

temperature of operation, which was controlled at a value in the neighborhood of 40 degrees C.

The method for frequency subdivision used in this first quartz crystal clock is illustrated in Fig. 11. The inductance element of an electric circuit oscillator, designed to operate at the desired low frequency, has a core of variable permeability so that the frequency can be adjusted over a narrow range through the control of direct current in an auxiliary winding. A harmonic of this low frequency, generated in the tube following the oscillator, is compared with the incoming high frequency in the vacuum tube modulator. The harmonic chosen has nominally the same frequency as that of the control, or crystal oscillator, so that one output of the modulator is a direct current whose magnitude and sign vary with the phase relation between the inputs to the modulator. The use of this method to regulate the low-frequency oscillator insures that the low frequency is some exact simple fraction of the high frequency. If, therefore, a synchronous motor is operated from the low frequency thus produced, its rate represents accurately that of the high-frequency source as though it had been possible to use that source directly.

Several other electrical circuits were proposed around 1927 for the subdivision of high frequencies. The method in most general use at present is an adaptation of the "multivibrator" first used by Henri Abraham and Eugene Block in 1919 for the measurement of high frequencies⁷¹. They used their circuit to produce a wave rich in harmonics and having a fundamental that could be compared directly with that of a tuning fork standard. By various means now well known the high frequency could be compared with one of the harmonics of this special oscillator.

This procedure was reversed by Hull and Clapp⁷², who discovered that the fundamental frequency could be *controlled* by coupling the high-frequency source directly into the circuit of the multivibrator. This, in fact, is a general property of any oscillator in which the operating cycle involves a non-linear current-voltage characteristic, being most pronounced in those of the relaxation type. Van der Pol and Van der Mark in 1927 reported on some experiments on "frequency demultiplication" using gas tube relaxation oscillators⁷³. The multivibrator is, in effect, a relatively stable relaxation oscillator⁷⁴, and with slight modification has been used extensively as the frequency-reducing element in quartz-controlled time and frequency standards throughout the world.

One serious difficulty with the multivibrator type of submultiple generator has been that, if the input fails or falls below a critical level, it will continue to deliver an output which, of course, will not then have the expected frequency. Certain variables in the circuit, such as tube aging, may cause a

similar result. With this in view, a general method for frequency conversion has been developed by R. L. Miller⁷⁵, in which the existence of an output depends directly on the presence of the control input. The basic idea involved in this, now known as regenerative modulation, was anticipated by J. W. Horton in 1919⁷⁶ but had not been developed prior to Miller's investigations. The circuit of a regenerative modulator in its simplest form as a frequency divider of ratio "two" is shown in Fig. 12.

Soon after the announcement in 1927 of the first quartz crystal controlled clock,⁷⁰ the idea was studied and applied in many places notably in America and Germany, and at the present time it forms the basis for precise measurements of time and frequency in many government physical laboratories as well as in many astronomical observatories and industrial and university laboratories throughout the world.

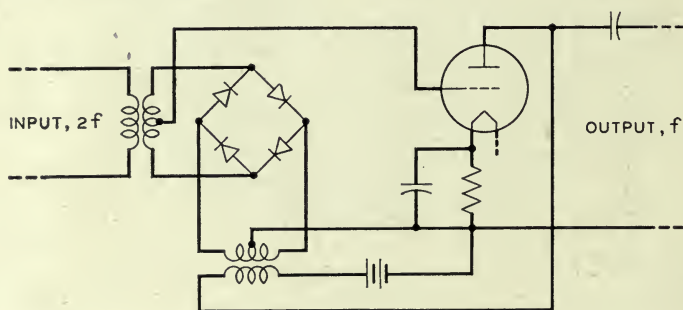


Fig. 12—Frequency divider for ratio TWO employing regenerative modulation.

Although the first results were quite satisfying, it was the immediate interest of all concerned to find out what improvements could be made, and these were not long in coming. As in the case of the pendulum already discussed, or with any other oscillator, the constancy of rate obtainable depends on two kinds of properties: those which concern the inherent stability of the governing device itself, and those concerned with the means for sustaining it in oscillation. Some of the factors in the two groups are interrelated and must be considered together.

The improvements in quartz oscillator stability therefore have been concerned with two main endeavors, namely that of cutting and mounting the resonator so as to realize effectively the unusually stable properties of quartz crystal itself, and that of coupling it to the electrical circuit in such a way as to avoid the effects of such variables as power voltage variation, aging of vacuum tubes, and the like, on the controlled frequency. The latter effects were not obvious at first because the temperature coefficient and the effects of friction and change of position in the mounting caused

variations of considerably larger magnitude. It was natural, then, to see what could be done about these effects.

Zero Temperature Coefficient of Frequency

With the knowledge that X-cut resonators had negative coefficients, frequently as large as thirty parts in a million per degree C, and that Y-cut resonators in general had positive coefficients, often in excess of a hundred parts in a million per degree, the author undertook to make resonators of such shape that the oscillations would occur in both modes simultaneously, and so combine the coefficients, in the hope that the resultant could be made zero.⁷⁷

The first experiments, made on two series of resonators both yielded encouraging results. The first was a series of rectangular X-cut plates of varying thickness shown in Fig. 13. The second was a series of three circular discs of different diameters, all being cut with the large surfaces in the plane of the Y and Z axes. The three discs were made from the *same* material, each smaller one being trepanned from the previous one after complete measurements had been made upon it. The set of circular crystals remaining after these tests were completed is shown in Fig. 14 and the slab from which they were cut is shown assembled with the original large crystal in Fig. 15.

Subsequent tests showed that the annular pieces could be designed for a low or zero coefficient and such a shape shown in Fig. 16 was employed for a number of years in the Bell System Frequency Standard in New York City⁷⁸. As described in this reference, the reason for using the ring in preference to the solid disc or rectangular plate was in the convenience of mounting. The rings were formed with a ridge in the central plane of the hole so that they could be supported on a horizontal pin thus providing a one-point support at a position where the vibration is very small. The rings used in this first application of zero coefficient quartz resonators have been called "doughnut" crystals for obvious reasons. In Fig. 17, George Hecht is shown making a final adjustment, by "lapping" with fine abrasive, on one of the four original zero-coefficient ring crystals. Mr. Hecht made all four of these resonators, as well as many others of various shapes and sizes used in the early experiments in this work.

Supported as described, the rings hang in a vertical plane and, as first used, they were supported freely between solid electrodes rather closely spaced to the flat surfaces. The small amount of free motion relative to the electrodes, inherent in this sort of mounting, caused occasional changes in frequency if the support were disturbed, which at times would be as large as one part in ten million. To avoid this difficulty, other ring crystals were

constructed with a sort of narrow shelf at the central plane that could be mounted in a horizontal plane on pin supports. The two methods of supporting the ring resonators are illustrated in Fig. 18. Such resonators were

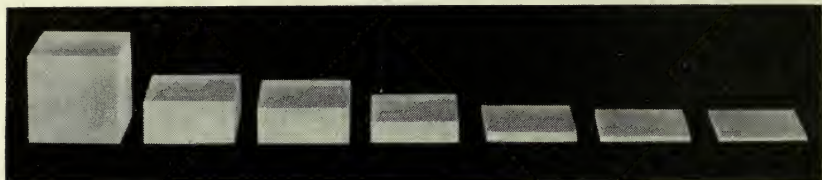


Fig. 13—Set of rectangular quartz resonators made for zero temperature coefficient study.

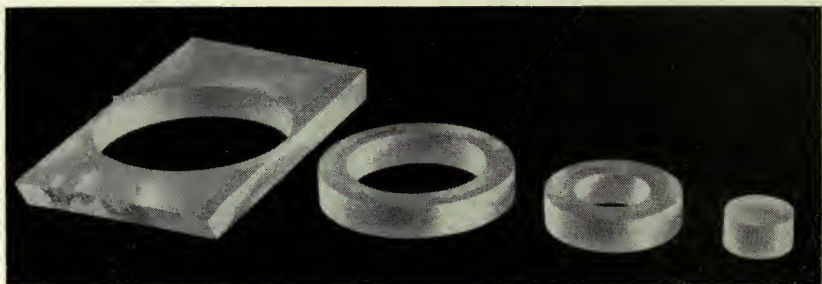


Fig. 14—Circular pieces remaining after temperature coefficient study of quartz discs and rings.

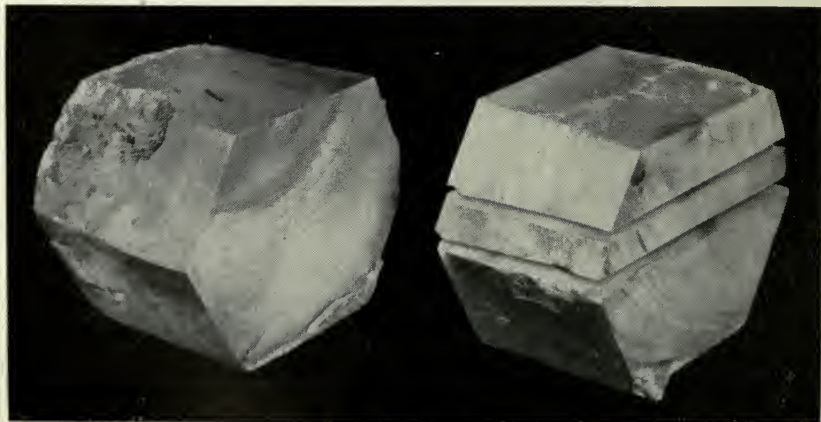


Fig. 15—Large crystal and slab from which low coefficient studies were made.

used in the Bell System Frequency Standard until 1937 when they were replaced by an entirely different type that will be described later.

The rings were adjusted to oscillate at 100,000 vibrations per second, the frequency which has been adopted in nearly all oscillators of extremely

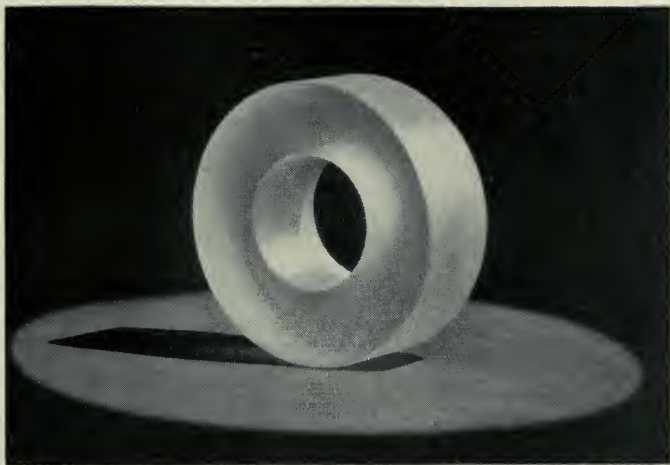


Fig. 16—100-Kilocycle quartz ring resonator with zero temperature coefficient.

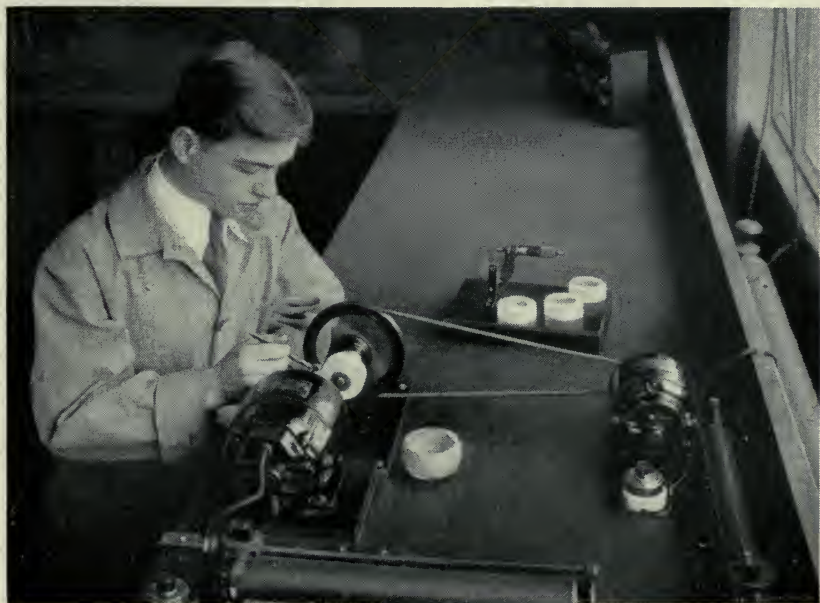


Fig. 17—George Hecht finishing the first set of zero-coefficient quartz rings.

constant rate. All of these rings were constructed to have a zero frequency-temperature coefficient at a temperature in the neighborhood of 40 degrees C, the frequency being a maximum at that point on an approximately para-

bolic characteristic. The zero temperature coefficient makes it possible to practically eliminate frequency changes caused by ambient temperature changes since, by relatively simple means, it is possible to control the resonator within ± 0.01 degree C, at the temperature for which the effect is substantially nil. The reduction of the effect of temperature, and the stabilization of the mounting, increased the stability of frequency control and oscillator-clock rate beyond anything that had ever been obtained before. Subsequent improvements that will be described later produced even greater stability.

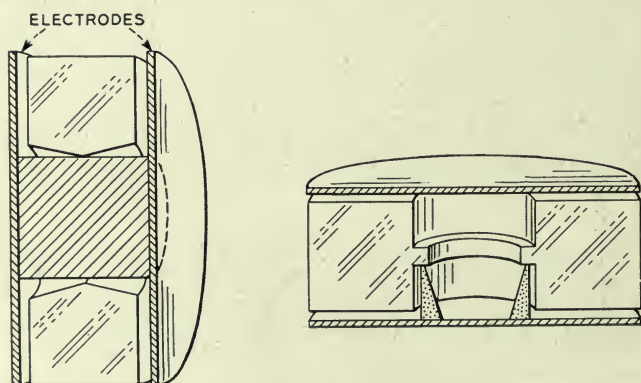


Fig. 18—Methods of mounting quartz ring resonators.

The Crystal Clock

The striking stability of the crystal oscillator clock led the author to propose the general use of this type of clock for precision timekeeping, the chief emphasis having been previously on the derivation of constant frequency. A paper entitled "The Crystal Clock,"⁷⁹ presented before the National Academy of Sciences in April, 1930, described such a clock and pointed out some of its properties and likely uses.

Chief among these properties, of course, is its inherent stability and relative freedom from extraneous effects. The quartz crystal clock is not dependent on gravity and, without any compensating adjustment, will operate at the same rate in any latitude and at any altitude. This property already has been useful in the *measurement* of gravity and gravity gradient by measuring the rates of pendulums on land and at sea.^{80, 81}

The crystal clock is practically immune to variations in level and shock and can be used as an instrument of precision under conditions entirely unsuitable to pendulum clocks. For this reason it performs satisfactorily in practically any location, including earthquake zones, and may be used in transit as in a submarine, in an airplane or on the railroad.

Some of the outstanding properties of the quartz oscillator clock were discussed in 1932 by A. L. Loomis and W. A. Marrison⁸², in relation to a series of experiments comparing the performance of quartz clocks at Bell Telephone Laboratories in New York and a set of synchronome free-pendulum clocks operating in The Loomis Laboratory in Tuxedo Park, about fifty miles away. The comparison was effected through a circuit maintained between the two laboratories over which a 1,000-cycle current controlled by a crystal in New York was used to drive the Loomis Chronograph⁸³ in Tuxedo Park. During part of the time, signals from the clocks were sent back over the same circuit and recorded on the Bell Laboratories' Spark Chronograph⁸⁴.

The quartz oscillator assembly at the Bell Telephone Laboratories at the time of these experiments is shown in Fig. 19. The four ring crystals in their individual temperature-controlled 'ovens' are mounted under hermetically sealed bell jars to avoid the effects of ambient temperature and atmospheric pressure changes. The vacuum tube oscillator circuits are immediately below the bell jars; and the control, monitoring and power supply equipment in the remainder of the space.

One of the most interesting results of these cooperative experiments was the measurement of a periodic variation in the rate of the pendulum clocks in phase with the lunar daily cycle. The amount of this daily variation is very small, being only a few tenths of a millisecond, but readily observable in comparison with a stable rate standard that does not vary with gravity.

Further Refinements in Quartz Clocks

The spectacular results from the use of the quartz crystal clock up to this time, about 1932, were due in part to its novelty and in part to the fact that it is quite independent of some of the variable factors that affect conventional precision clocks, including gravity itself upon which the rate of all pendulum clocks depends. The remarkable stability of present day quartz oscillators and clocks is the result of a series of developments and refinements extending over a number of years.

As mentioned previously, the factors that cause departure from constant rate in the completed operating device fall into two distinct classes, namely those which concern the inherent or natural frequency of the resonator itself, and those which concern the means for driving it at that inherent rate.

The first class comprises all those properties of the mounted resonator which tend to relate its inherent rate to ambient conditions such as temperature, atmospheric pressure, change of position and vibration, and to the passage of time—that is, aging. Since the final stability cannot exceed the inherent stability of the mounted resonator itself, its study is of prime importance.



Fig. 19—Bell System Frequency Standard, 1930.

The second class comprises properties of the means for sustaining oscillations in such a resonator which relate the resulting actual rate to variations in the electrical circuits, in the power voltages, in vacuum tubes and other like effects. In the limit, it is the hope that the net result of all such effects can be eliminated so that the stability of the quartz crystal alone will remain the sole governing factor. This is the goal, and the inherent stability of the substance, quartz crystal, is the limit toward which the stability of the quartz crystal clock will approach but cannot exceed.

The development of the quartz resonator and its mounting for numerous applications is described in some detail by Raymond A. Heising and his collaborators⁸⁵ in their recent book, "Quartz Crystals for Electrical Circuits". Of all the types of resonator described in this work the one having the most extensive use at the present time, for quartz clock installations and for other applications of comparable accuracy, is the GT crystal resonator developed by W. P. Mason⁸⁶. This resonator is cut from quartz crystal in such a way that the positive and negative coefficients are effectively neutralized over a range of about 100 degrees C, so that in any part of this range the resulting temperature coefficient of frequency is not more than one part in a million per degree C. With suitable precautions in manufacture, the tangent at the point of inflection in the frequency-temperature curve may be made horizontal, which means that the temperature coefficient may be made substantially zero over a considerable range of temperature.

The GT crystal resonator therefore introduces two significant advantages in timekeeping, namely that greater accuracy of rate may be obtained with a given accuracy of temperature control and that the value at which the temperature is controlled may be chosen in a considerable range. In fact, without any temperature control at all, the rate of a clock regulated by such a crystal may be accurate to a tenth of a second a day over an ambient range of 100 degrees C. Among the many quartz clock installations now using the GT resonator, all or in part, are the Royal Observatory at Greenwich, the British Post Office, the U. S. Naval Observatory and the U. S. Bureau of Standards.

One of the chief sources of variation in rate of quartz oscillators, in the early stages of their development, was in the means for mounting and in the electrical circuit connections. As mentioned previously, any variation in the effective resistance or in the effective mass or stiffness of a resonator has a direct effect upon its rate of oscillation. The problem reduces to that of supporting the resonator so that the frictional losses are small and constant and so that the coupling to the electrical circuit is as nearly as possible invariable.

The mounting of quartz crystal units is discussed at length by R. M. C.

Greenidge in Chapter XIII of Mr. Heising's book referred to above.⁸⁵ The most satisfactory means by far that has been found for mounting crystals of the GT type is that of actually soldering them to thin supporting wires by means of small discs of silver deposited on the crystal at its nodes. This method serves the double rôle of supporting the crystal and of providing electrical connection to metal electrodes plated on the crystal. Resonators so supported may be made almost immune to mechanical shock and will continue in satisfactory operation through accelerations of several times g . Nearly all crystals which vibrate in a long dimension are now mounted in this way. One manufacturer produced about 10,000,000 crystals of a single

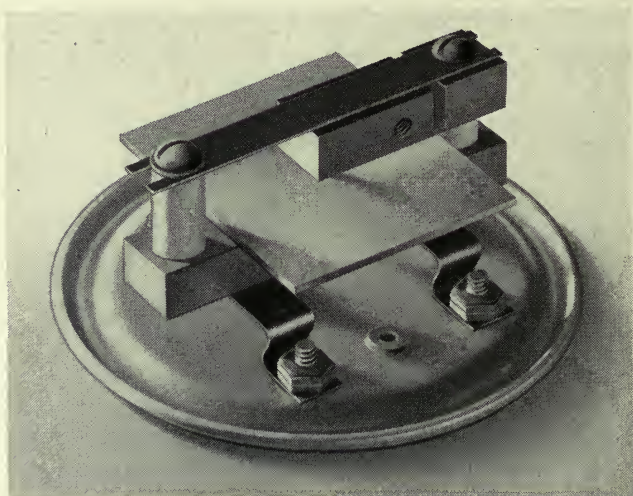


Fig. 20—Pressure-mounted GT crystal for sealing in a metal envelope.

type so mounted in a three-year period during World War II. Prior to the use of wire supports, such crystals were "pressure mounted" by means of small metal jaws which clamped from opposite sides at the nodes. A GT crystal mounted in this way is shown in Fig. 20. Crystals so mounted are still in use in the Bell System Frequency Standard, being the first of the GT crystals to go into actual service. This type of mounting is not quite so stable as the wire mounting and is somewhat more difficult to manufacture. One of the wire-mounted crystals such as developed for LORAN and other oscillators of comparable accuracy is shown in Fig. 21.

The plating of electrodes on the crystal surface has led to increased stability of frequency control, chiefly because the coupling to the electrical circuit may be kept more nearly constant thereby. When separate electrodes were employed, the variation in spacing was always found to be a

source of instability, as mentioned previously in relation to the use of the first ring crystals. Plating of crystals is not a new idea but the application to quartz resonators of high Q requires a great amount of technical skill in order to obtain coatings which are mechanically and chemically stable and which utilize the minimum of added material. The use of too much metal

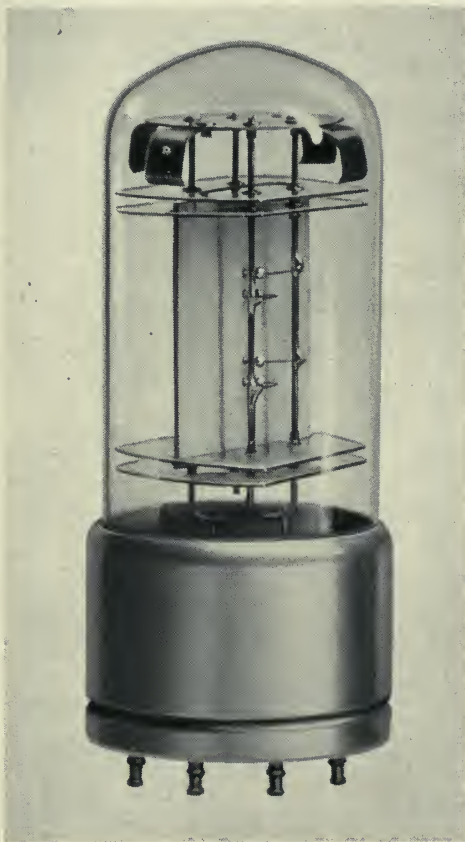


Fig. 21—Wire-supported GT crystal sealed in a glass envelope.

will, of course, impair the resonator by increasing its rate of energy dissipation and probably its aging rate. The metal most often used for electrodes is silver, although gold and aluminum have been used in special cases. Evaporation in vacuum has been found to be the most satisfactory method for the actual plating, giving very adherent coatings and being subject to precise manufacturing control. The art of plating quartz resonators is discussed in detail by H. W. Weinhart and H. G. Wehe in Mr. Heising's book.

Several other factors have had an important bearing on the final stability of quartz resonators. One of the most important of these is the care that must be exercised during fabrication in order to avoid setting up stresses in the material that subsequently can be relieved only slowly. By slow grinding with adequately fine abrasive such effects can be kept very small. Etching with hydrofluoric acid has resulted in much further improvement through the removal of stressed surface material and all potentially loose material which, formerly, often caused anomalous aging effects. Artificial aging by heating, and thorough cleaning before and after plating, have also contributed greatly to the final stability of the crystal unit. The resonator finally is mounted in high vacuum in a glass envelope in order to eliminate losses due to sound radiation and friction, and to protect it from surface contamination and chemical action.

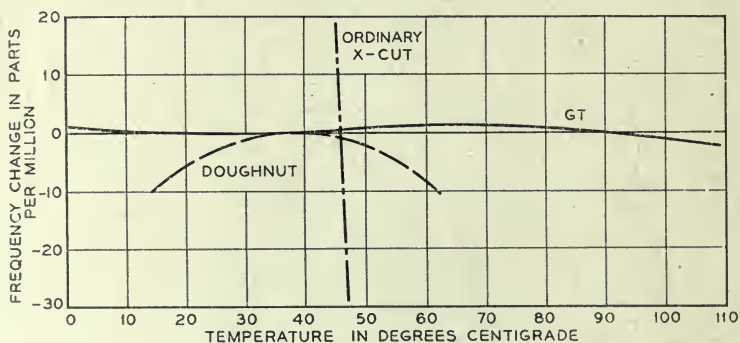


Fig. 22—Frequency-temperature characteristics for three types of quartz resonators.

Even the most perfect quartz resonator, in an ideal mounting, is unable to keep time unless it is maintained in oscillation; and, like a pendulum, its rate will depend in large part on the manner in which it is driven. The same general principles apply to both cases, except that usually a pendulum is driven by impulses which should be applied when the velocity is maximum, while a quartz resonator is usually driven by a sinusoidal force arising through the piezoelectric coupling, and so phased that the maximum force occurs when the velocity is maximum. This, in fact, is a required condition for maximum rate stability. The graphical analysis of Fig. 5 applies equally for the case of sine wave drive, since the sine wave can be considered as the summation of an impulse at its peak and of sets of pairs of impulses symmetrically disposed with respect to it. Obviously, the phase errors for each such pair of impulses cancel, bringing us back to Airy's condition, but with the broader view that, for the feedback or driving wave to have minimum effect on the rate of an oscillator, the force wave must be in phase with the velocity of the resonator.

Numerous vacuum tube circuits have been proposed and used for maintaining quartz resonators in oscillation, some of which are illustrated in Fig. 8. The one among these which at present most nearly approaches the ideal is that developed by L. A. Meacham, known as the Bridge Stabilized Oscillator.⁸⁷ This oscillator, in its original form or with slight modifications, is now used almost universally in England and America where the maximum stability of rate control is required.

In the bridge stabilized oscillator, the feedback path is through a Wheatstone bridge with the crystal in one arm and with resistances in the other three. The frequency of oscillation becomes that for which the reactance of the crystal approaches zero; the bridge can only be balanced when the crystal behaves electrically like a resistance. The unbalance voltage from the bridge is fed back into the amplifier, which should provide a relatively high gain, as will appear. The great frequency stability of this oscillator depends upon the fact that, in the neighborhood of balance, a small phase shift in the resonant elements causes an enormously larger phase shift in the unbalance voltage. But the actual amount of this unbalance phase shift is limited by the fact that it must be equal and opposite to that in the amplifier in order for oscillations to be sustained. This insures that at all times the phase shift in the crystal is much smaller than that occurring in the amplifier which itself can be made small by suitable design. The ratio of the phase shift of the bridge output to that of its input increases as balance is approached, making it possible to practically eliminate the effect of phase shift in the amplifier simply by increasing the amplifier gain. Most of the variable factors in the amplifier of an oscillator circuit affect the controlled frequency through the phase shifts caused by them. It is evident, then, that the bridge circuit, which permits only a small fraction of such phase shifts to become effective at the resonant element, will substantially free the resonator from variable effects in the amplifier and allow it to control a rate determined almost wholly by its own properties.

When the above condition is attained and the crystal resonator, when oscillating, acts in the circuit like an electrical resistance, it acts that way *because* the velocity is in phase with the applied mechanical force, which, as has been stated, is the condition for most stable rate control. In the crystal oscillator, this ideal condition is obtained simply by the automatic balancing of a bridge circuit, accomplishing in a most elegant manner the equivalent, in the case of a pendulum, of applying driving pulses at the exact center of swing.

The bridge-stabilized oscillator includes also an automatic control of amplitude. The variation of frequency with amplitude is very small and in no way comparable with the "circular error" of an ordinary pendulum, but in the quest for the highest attainable stability it must be taken into account.

The control of amplitude is obtained by the use of a resistance with positive temperature coefficient in the bridge arm conjugate to the crystal, chosen so as to have exactly the right value to balance the resistance of the crystal when a specified current is flowing in the bridge. If larger than normal current flows momentarily the resistance is increased, which decreases the feedback, thus stabilizing the amplitude at some predetermined value. For the highest stability it has been found advantageous to operate the crystal at a very small fraction of the amplitude that normally would be used in a power oscillator. In power oscillators the crystal sometimes is subjected to strains near the fracture point, which is not a favorable condition for precision control. The actual amplitude of motion of the crystal is of course extremely small. In the GT crystal, as currently used, the maximum change of dimensions during oscillation amounts to only about ± 0.0006 per cent.

The improvements in quartz resonators, and in their driving circuits, have resulted in the construction of quartz crystal clocks that will keep time with an accuracy better than 0.001 second a day, so that measurements of time of great interest and value to astronomers and geophysicists can now be made with an accuracy hitherto unattainable.

Facility of Precise Time Measurement

In making such precise measurements of time it is of importance, second only to the inherent accuracy of the standards themselves, to have available means whereby they can be carried out with facility and within a reasonable time interval. The ease with which precise time measurements, and precise rate comparisons, can be made is an outstanding feature of the quartz crystal clock and already has an important bearing on the use of this type of clock in astronomical observatories. This facility depends chiefly on two properties of the oscillator clock: first, that *continuous* rotation of controlling and measuring devices can be produced having the stability of the primary control element; and, second, that the period of the control element, and therefore of alternating current controlled by it, is of very short duration.

The first of these, through simple devices controlled directly from the electrical output of the crystal oscillator, with suitable frequency reducing equipment, permits of ready comparison between any time phenomena in the form of electric or light signals, and of the derivation of precisely controlled time signals for radio transmission and for laboratory experiments.

Of prime importance among these comes the means for rating crystal clocks in terms of stellar observations using meridian transits or the photographic zenith tube⁸⁸. It is possible to control a mechanism in the time-star observing equipment so that the difference between a star position *predicted* from the clock rate, and the *actual* star position, can be observed directly or re-

coded photographically with great accuracy. The difference thus observed, after allowing as well as possible for known systematic errors, is the best known single check on the time indication of a clock. A series of such observations constitutes the best known measure of the *rate* of a clock. The great value of the method is that the comparisons are made directly without the need of any intermediate mechanism thus eliminating a large part of the "personal error" of observation. The probable error of observation as derived from a number of such measurements on a good night may be as small as one or two milliseconds⁸⁹. The average rate of a clock thus determined depends on the number of days over which the rate is computed and in a two-week period may be compared with the rate of the earth, that is, with astronomical time, with an accuracy of one part in one hundred

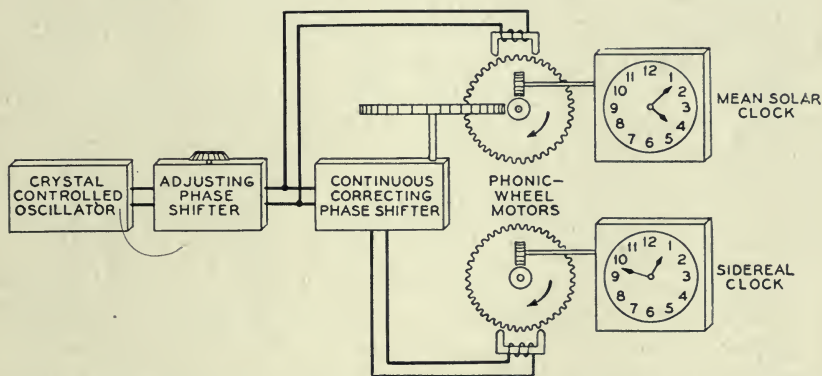


Fig. 23—The use of an electrical phase shifter to adjust the timing of a signal. (From "The Crystal Clock", 1930)

million or about a third of a second a year. All this, of course, is contingent on the stability of the quartz clock, which, except for long-time effects, may be demonstrated independently.

A rotating mechanism controlled directly from a crystal clock is admirably adaptable to the transmission of precise time signals. Rhythmic signals of any desired structure can be produced readily by means of cams, special generators, or interrupted light beams, and the timing of those signals can be adjusted as precisely as the clock time is known by simply advancing or retarding the signal generators. Such adjustment is attained readily by means of differential gearing in the mechanical system, or by means of continuous phase shifters in the electrical driving circuit. The use of electrical phase shifters for this purpose was first proposed in "The Crystal Clock" paper⁷⁹ previously mentioned. Figure 23, taken from that paper, illustrates the manner of using the phase shifter with one type of time signal

generator. Extremely fine control of timing is possible by means of the electrical phase shifter since it can be included in the circuit at any stage of frequency subdivision. If, for example, it is used at the lowest frequency, assumed to be 1,000 cycles, one complete turn of the phase shifter dial will cause a progressive time adjustment of one millisecond. When used at a higher frequency, the precision of adjustment is increased correspondingly. Continuous phase shifters suitable for such purposes were proposed as early as 1925.⁹⁰ The idea of utilizing continuous phase shifters for the purpose of making controllable changes in the frequency or indicated time in a standard time and frequency system⁹¹ was first disclosed in a comprehensive patent filed in 1934 and issued to Warren A. Marrison in 1937. The most elegant type of phase shifting element suitable for such purposes was developed by Larned A. Meacham.⁹² This has been used in many transmission systems requiring continuous variation of phase such as in variable direction radio beam systems⁹³ and LORAN.

The conversion between mean solar time and sidereal time, or for that matter between any time systems, may be accomplished very easily with the quartz clock. Having a rotating device, such as a dial or commutator, whose rate corresponds to mean solar time, it is only necessary to apply a gearing or the equivalent to obtain another rate corresponding to sidereal time. It has been shown by F. Hope-Jones⁹⁴, Ernest Esclangon⁹⁵ and others how any desired ratio, such as the ratio of the rates of mean solar and sidereal clocks, can be obtained with any required accuracy by gearing. A combined mechanical and electrical method was proposed in the "Crystal Clock" paper by means of which this ratio can be realized with an accuracy of one part in 10^{11} using simple gearing and a continuous phase shifter.

The potential value of the factors just discussed in precision time studies was realized early in the crystal clock development. This was indicated in the "Crystal Clock" paper written in 1930 which closed with the following paragraph:

"It would thus be possible to combine, in a single system mean solar and sidereal time-indicating mechanisms, means for rating the clocks in terms of time star observations and means for transmitting time and frequency signals with the absolute accuracy of the time determinations."

It is of some interest to compare this prediction with the present trend of development. In describing the quartz clock installation at the Royal Observatory in Greenwich, Sir Harold Spencer Jones stated⁸⁹ in 1945:

"The quartz clocks being installed at the Royal Observatory are all adjusted to give a frequency of approximately 100,000 per mean time second. By suitable gearing, the synchronous motor can give impulses every sidereal second and tenths of seconds. Thus, the same clock can be made to serve both as a mean time and as a sidereal time standard. All time signals are, of course, sent out according to mean time; the sidereal time is required only for the actual time determination so that it is not necessary for all the clocks to have the gearing to give sidereal seconds."

The importance of the convenient methods for measuring time and time interval inherent to the crystal clock is emphasized by the fact that some observatories employed crystal clock mechanisms in connection with stellar observations and in the transmission of time signals before they were used in the actual time *keeping* department⁸⁸.

The second property contributing greatly to the convenience of precise time measurements is the relatively very short period of the quartz clock control element. The chief advantage lies in the extreme accuracy with which the rates and indicated times can be compared by electrical methods. An example will suffice to illustrate this point.

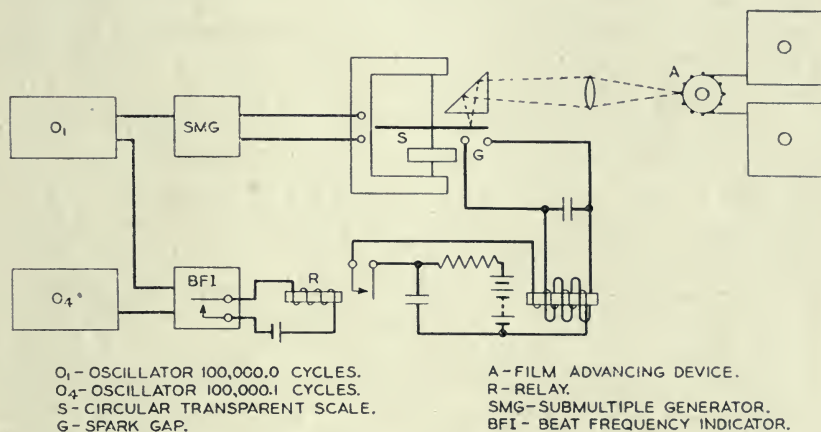


Fig. 24—Device for comparison of oscillator rates accurate to 1 part in 10,000,000,000. (From "High Precision Standard of Frequency", 1929.)

Since the *rate* of a crystal clock is the rate of oscillation of the crystal or of the current driving it, it is only necessary, in comparing clock rates, to measure the relative frequencies of the oscillators concerned. This can be done by any of the standard methods for frequency comparison⁹⁶ but, in the case of quartz clocks, since in general the primary frequencies are high and are nominally the same, special methods of extreme accuracy can be employed. The apparatus first designed for the ultra-precise comparison of quartz oscillators and capable of an accuracy of one part in 10^{10} was described by Marrison in 1929.^{78, 97, 98} The principle of its operation is shown in Fig. 24, reproduced from the paper "High Precision Standard of Frequency".

Two oscillators to be compared were adjusted so as to differ by about one cycle in ten seconds. The problem reduces to that of measuring the beat frequency, nominally 0.1 cycle per second, with as great accuracy as possible.

This was done by measuring the duration of each beat by a photographic method. By means of a modulator, a relay, and induction coil, a spark was produced at the spark gap at a definite phase of each beat period. The spark illuminated the edge of a transparent scale rotating 10 revolutions per second under control of one of the oscillators, O_1 . The transparent scale contained 100 numbered divisions, which therefore represented milliseconds in any time interval so measured. Each time a spark occurred, the portion of scale illuminated was registered on photographic film. Thus, the duration of each beat was registered photographically with an accuracy of one part in ten thousand. Since the beat frequency is one millionth of the high frequency, the resulting *comparison of high frequencies* is precise to one part in ten thousand million, or 1 in 10^{10} . Actually, it was possible to estimate fractions of a scale division which gave greater precision of measurement than was required in the study of oscillators of that date.

L. A. Meacham in 1940 improved upon this method of frequency comparison by substituting an electronic relay for the mechanical relay, and by using a discharge lamp instead of a spark for illumination. He used the improved apparatus⁹⁹ for studying the behavior of the then new and highly stable bridge stabilized oscillators.

Still further improvements in the general method have been reported by H. B. Law using a "phase discriminator" to trigger off a special chronometer, consisting of a decimal scaling counter, and thus avoiding the photographic process¹⁰⁰. The scaling counter as used here counts the number of cycles of a 100,000-cycle input timing wave that occur during any one beat between the two frequencies being compared, and registers that number, in scale of ten, on a system of dials that can be read directly. In comparing frequencies that are free from interference, the accuracy of comparison by this means is limited chiefly by the precision with which the "phase discriminator" can mark the beginning of successive beats. An accuracy of one part in 10^{11} is claimed. This is one of the rate comparison means employed in the frequency and time standards of the British Post Office and in measurements involving the quartz clocks of Greenwich Observatory and the National Physical Laboratory.

The scaling counter is a particularly useful device for the precise measurement of any time or rate phenomena that can be reduced to the measurement of short time intervals. The counter idea originated some years ago as a means for counting alpha-particles and other phenomena associated with radioactivity studies, one of the original devices being the well known Geiger-Muller counter. The basic scaling circuit, used in many counters, was proposed in 1919 by W. H. Eccles and F. W. Jordan. An interesting history of counting circuits as applied primarily to the counting of electron and nuclear particles has been written by Serge A. Korff in his book on that

subject published in 1946.¹⁰¹ The early scaling circuits operated on the binary system, but recently various circuits have been developed that give the count in scale-of-ten notation with certain advantages, chiefly that of convenience, associated with the common decimal system of notation. A discussion of some modern binary and decade electronic counters¹⁰² was published by I. E. Grosdoff in September, 1946.

Methods of measurement such as this, and the stable properties of the quartz clock which make them desirable, are of importance in the precise measurement of time because the *nature* of variations in rate, so small that, if continued unchanged they would accumulate to only one second in a thousand years, may be studied under controlled conditions in the laboratory, and with such facility that a comparison with this precision can be made every ten seconds.

In a simpler manner, the short period of one oscillation of the quartz oscillator is of direct interest to the astronomer in connection with means for the intercomparison of his clocks in time. This reduces simply to counting the number of cycles gained or lost by one oscillator, referred to another, and may be accomplished in a great number of ways, yielding, on the basis of whole numbers of cycles, an absolute accuracy of time comparison of 0.00001 second.

An elegant method for accomplishing this¹⁰³, which also indicates automatically which clock is fast or slow, employs a special vacuum tube circuit to produce a polyphase current having the frequency *however small* of the difference between any two oscillators nominally the same. This polyphase current is used to operate a special synchronous motor whose angular position corresponds at all times to the phase angle of the vector representing the polyphase current. This relation holds all the way to zero frequency difference, in which condition the angular position of the motor, now at rest, indicates the phase relation between the two high frequencies. If the beat frequency goes through zero, the motor reverses. By this means, it is possible with very simple equipment to set up dial indicators showing continuously the time comparisons between any group of quartz clocks, taken in pairs, with an absolute accuracy of 0.00001 second. Of course, to operate other indicators, contacts, etc. from this device is a simple mechanical problem.

The principle of operation of the polyphase modulator is illustrated in Fig. 25, which shows one of the many possible forms of this device. Other modulator elements than vacuum tubes are used in some applications. In the form shown here it is necessary only to assume that the vacuum tubes produce second-order modulation, the lowest-frequency component of which is employed. If inputs at the two frequencies f_1 and f_2 , which are nearly the same, are delivered into the two balanced modulators *A* and *B* in such a

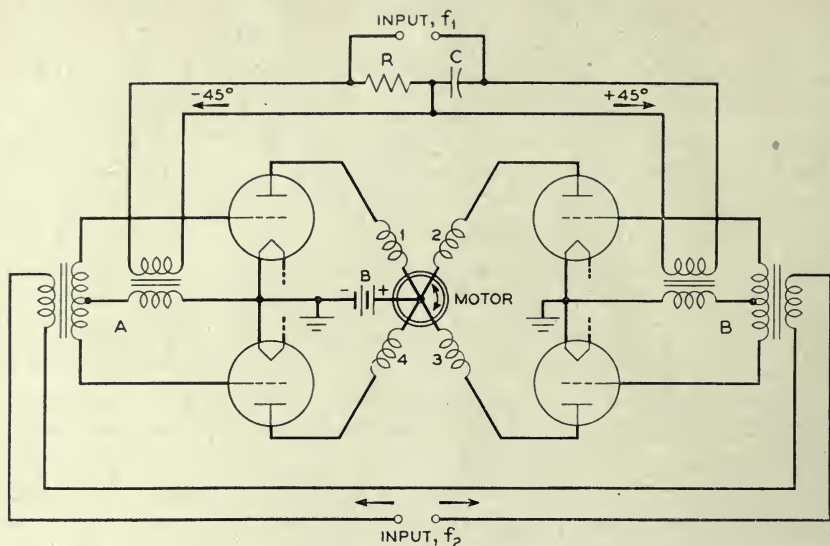


Fig. 25—Polyphase modulator for the absolute comparison of two oscillators of nearly the same frequency.

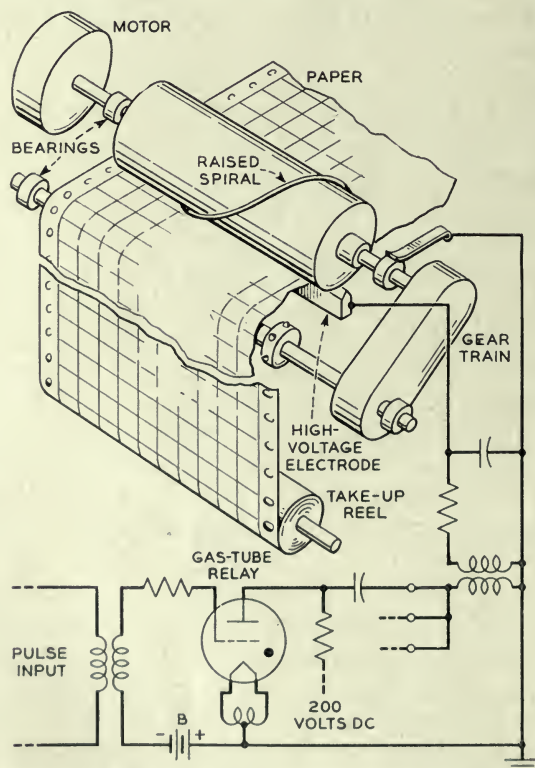


Fig. 26—Spark chronograph—schematic of operation.

way that there is a 90-degree phase shift between the two input voltages for one of the frequencies, the lowest-frequency component appears as a sinusoidal current in the output circuits 1, 2, 3 and 4 separated in phase by 90 electrical degrees in cyclic rotation. The principal output, therefore, is a 4-phase current having the frequency of the difference between the two inputs. If the magnetic circuits are arranged geometrically as shown, the resulting

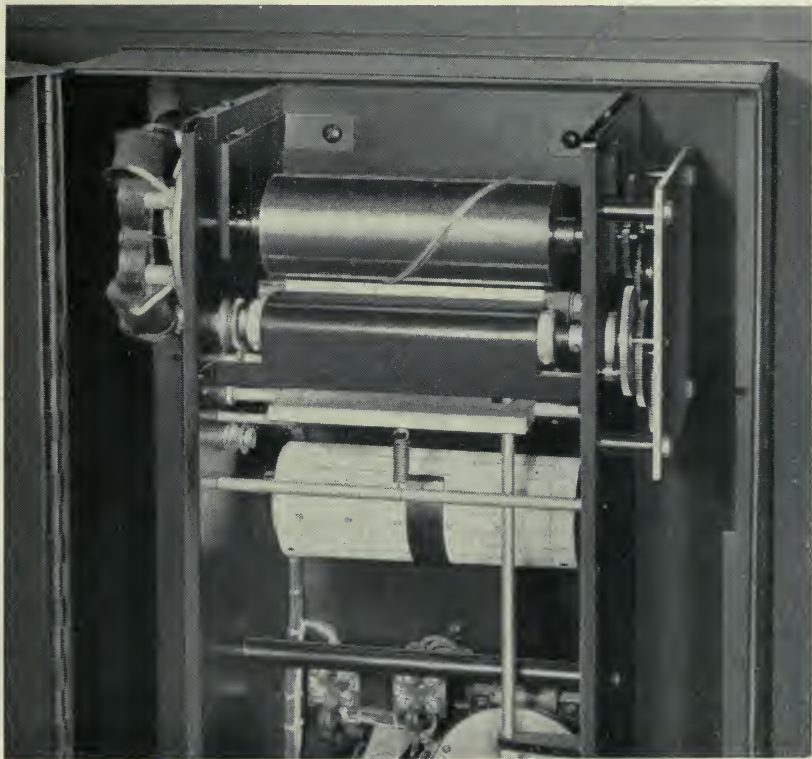


Fig. 27—Spark chronograph—close view of mechanism.

magnetic vector will rotate clockwise or counterclockwise depending on which frequency is high, or will remain stationary, indicating the phase relation, if the two frequencies are exactly equal.

Motors have been designed and are commercially available suitable for operating synchronously from such polyphase modulators, and form an excellent basis for the intercomparison of quartz oscillators and clocks with ultra-high precision.

For making records of time comparisons the spark chronograph⁸⁴ shown in Figs. 26 and 27 has served a very useful purpose, combining in a single

convenient instrument the means for comparing recurrent time phenomena with an accuracy of a millisecond or two on a continuous chart which shows the records for an entire week. Electrical impulses, related to the time phenomena to be recorded, operate trigger tubes which discharge condensers through the primary of an induction coil and cause sparks to jump from a rotating spiral through a special chart paper having a dark colored backing

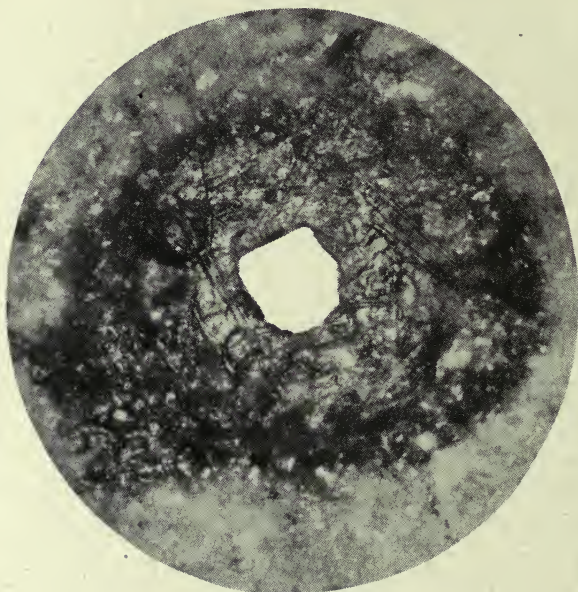


Fig. 28—Photomicrograph of single spark record showing nature of recording on wax-coated chart paper. $\times 100$

and coated with a very thin layer of white wax. As the chart paper moves slowly under the spiral, corresponding to the time abscissa, the succession of sparks produces readily visible traces consisting of rows of tiny holes with small areas around them where the wax is melted revealing the dark background. The holes are so small as to be scarcely visible, the darkened areas constituting the visible trace. Figure 28 shows an enlargement of the record of a single spark illustrating the nature of the marking. A recorder¹⁰⁴ very much like the Bell Laboratories' spark chronograph is used currently as part of the standard frequency and time broadcast equipment of the U. S. Bureau of Standards.

APPLICATIONS OF QUARTZ CLOCKS

The many useful properties of the quartz crystal clock have been the reason for its wide and expanding application for the precise measurement of time and rate.

First in historical order was the application to the measurement and control of frequency in communication. In this, the clock, through comparisons with astronomical time,⁶ served as the means for determining the frequency controlling it, the stability from the outset being great enough over intervals of a day or more so that the average rate, as determined by daily checks with time signals, was a very close approximation to the instantaneous rate at any time intervening. The first of these clocks, already referred to⁷⁰, was constructed in 1927 at the Bell Telephone Laboratories, in New York City, primarily for use as an accurate standard of frequency. Since that first experiment, three subsequent installations have been built in replacement with progressively improved performance. The standard now in operation (1947) was installed in 1937, using the first laboratory model GT crystals and the first set of four bridge-stabilized oscillators, and has been in operation continuously since that time. Two of the four oscillators, mounted in a temperature controlled booth, are shown in Fig. 29, and part of the auxiliary equipment, including a clock dial, a spark chronograph and some monitoring equipment, is shown in Fig. 30. This apparatus serves as the standard for precise measurements of frequency and time throughout the Bell System and is used to regulate the telephone Time of Day Service in New York City. It is the standard of reference for the electric light and power services in Metropolitan New York¹⁰⁵, and is used for a number of other similar services, distributed through the medium of a submaster installation¹⁰⁶ maintained by the Long Lines Department of the American Telephone and Telegraph Company. The original oscillators in this submaster installation were controlled by electrostatically-coupled 4000-cycle steel tuning forks *in vacuo* but recently have been replaced by improved oscillators controlled by 4000-cycle bi-morph quartz resonators.

A clock shown in Fig. 31, which is on display in a window of the American Telephone and Telegraph Company at 195 Broadway, is controlled from this source. It is sometimes called "The World's Most Accurate Public Clock".

The facility with which standard frequency and time services can be provided and distributed is an outstanding feature of the quartz clock development. Such services, having the accuracy of the primary controlling standard, may be provided anywhere that can be reached through a suitable

communication channel. As an example of this, a new primary standard equipment is being constructed for installation at the Murray Hill, New Jersey location of Bell Telephone Laboratories, the services of which will be

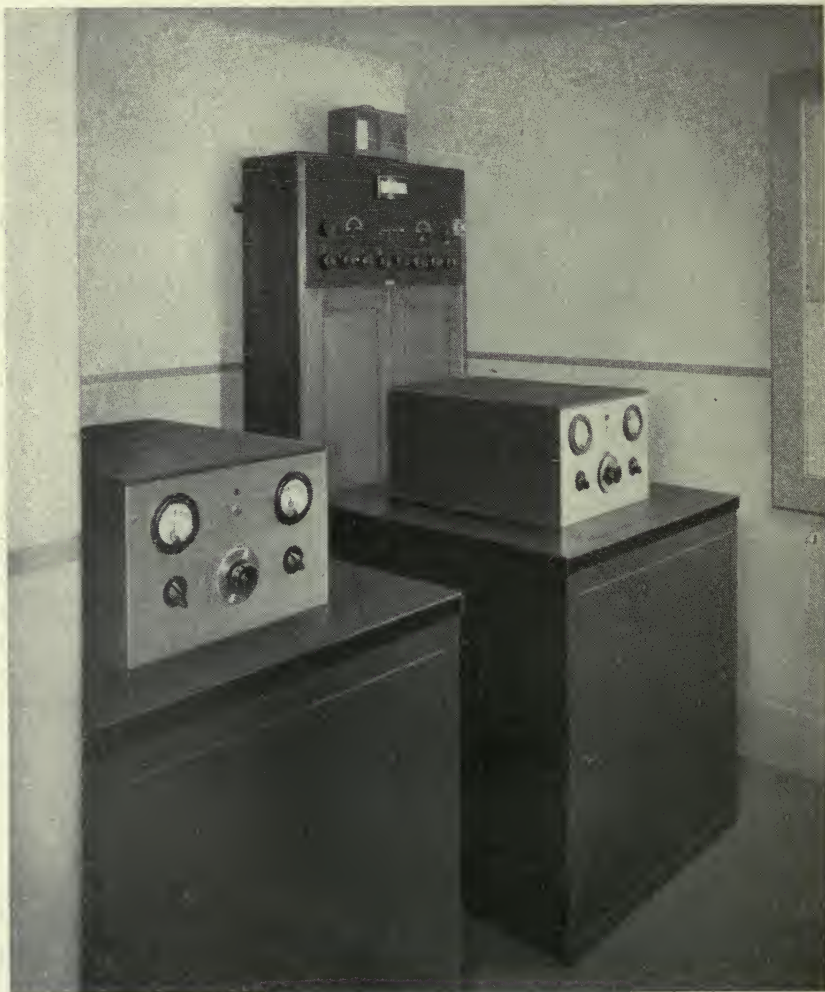


Fig. 29—Two of the four quartz oscillators of the Bell System Frequency Standard, 1937 to date.

made available through permanant wiring to all departments concerned. A number of frequencies in the range from 60 to 10,000,000 cycles, all controlled from the same crystal source, will be made available at some thirty locations at the Murray Hill Laboratories, as well as to other laboratories

of the Bell System and, through the Long Lines Department, to outside agencies.

A considerable number of quartz clocks have been built and used in laboratories and observatories all over the world, some as standards of



Fig. 30—Clock dial and monitoring equipment associated with the Bell System Frequency Standard, 1937 to date.

frequency, some as precise clocks, and others for general use in all measurements of rate and time. It would be impossible to mention all of these, for already there are many of them. But certain installations are of especial interest and will be discussed briefly.

When the Crystal Clock was first described as such in April 1930, the idea was discussed quite widely in Europe and America, and it was not long before the work was duplicated and extended in other places. The first outstanding application of the quartz clock to astronomy was made in

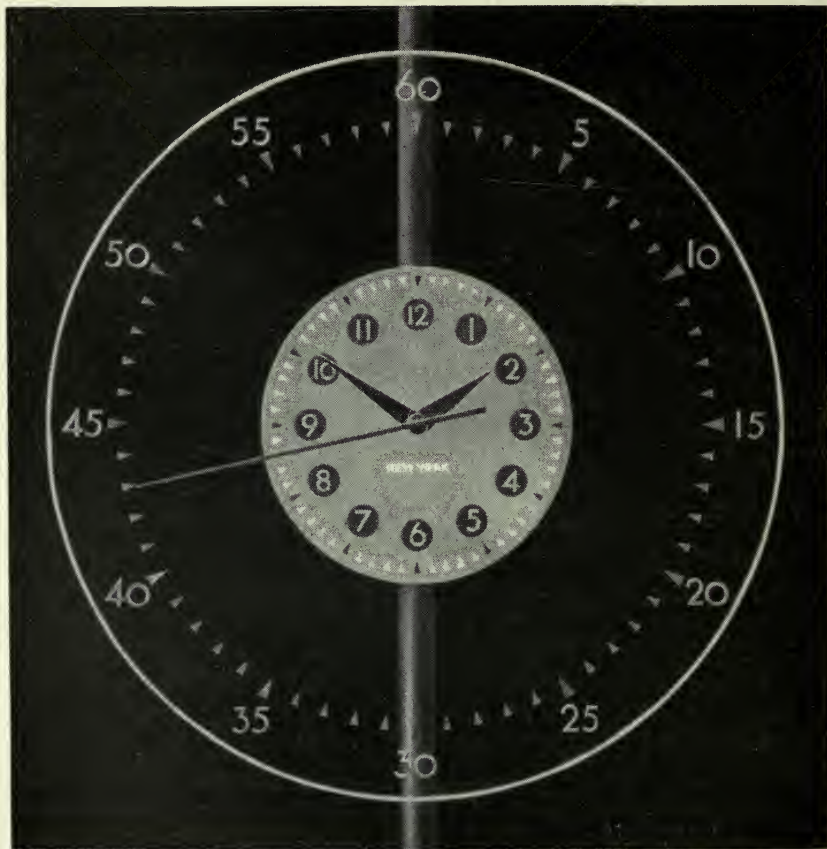


Fig 31—Display clock at 195 Broadway, New York. This clock, controlled by the Bell System Frequency Standard, shows the same time as that of the New York Telephone Time Service.

Germany with the installation at the Physikalisch-Technische Reichsanstalt. This was described by Scheibe and Adelsberger in 1932¹⁰⁷ and 1934¹⁰⁸, and reports of its splendid performance continued periodically. It was with this installation that it was possible for the first time to observe and measure variations in the earth's rate occurring over intervals as short as a few weeks. Previous measurements of such variations, involving studies of motion of the moon, the planets, and Jupiter's satellites, had required years to obtain

comparable information which, of course, by nature, could never reveal short-term factors.

Soon after the inauguration of the quartz clocks at the Physikalisch-Technische Reichsanstalt, somewhat similar installations were made at the Prussian Geodetic Institute at Potsdam¹⁰⁹, and at the Deutsche Seewarte in Hamburg¹¹⁰. The latter has been moved because of war conditions and is now the Deutsche Hydrographische Institut. The quartz resonators used in these installations are believed to be similar to those in Clocks III and IV in the Physikalisch-Technische Reichsanstalt installation except that some of them were made for 100 kilocycles instead of the original 60 kilocycles. They were made by the firm Rohde and Schwarz where also is maintained a quartz clock installation of extremely high precision¹¹¹.

For a number of years the U. S. Bureau of Standards at Washington, D. C. has maintained a quartz clock installation for their extensive constant frequency and time services. The early history of this installation was described in some detail by E. L. Hall, V. E. Heaton and E. G. Clapham in 1935.¹¹² As is now well known, the Bureau broadcasts a number of precisely controlled carrier frequencies at all times, all of which carry standard time and frequency modulations, including audible pitch standards and time signals. The audible pitch standards are 4000 cycles and 440 cycles, while the time signals consist of a succession of seconds pulses, continuous except for certain omissions for the purpose of identifying longer time intervals. All of these rates, including the carrier frequencies, are derived directly from crystal oscillators and are known so well that their accuracy as transmitted is estimated as one part in 50,000,000 at all times. The relative rates of the standard oscillators are compared and recorded continuously at the Bureau of Standards with an accuracy of one part in 10^9 . The time signals involved in these transmissions are so precise, and so convenient to use, that they may be employed for the high-precision intercomparison of quartz clocks across the Atlantic and for studies in astronomical time, heretofore difficult or impossible to accomplish by any other means.

The present standard frequency and time service facilities at the U. S. Bureau of Standards, which have been instituted under the general direction of J. H. Dellinger, are described in recent separate articles^{113, 114} by Vincent E. Heaton and W. D. George respectively of the Bureau, both of whom have made very substantial contributions to this development. The transmitting station for the standard frequency broadcasts, which comprises a complete set of quartz oscillators and control and measuring equipment, is shown in Fig. 32.

The absolute rates for the crystal oscillators at the Bureau of Standards are determined through cooperation with the U. S. Naval Observatory, also

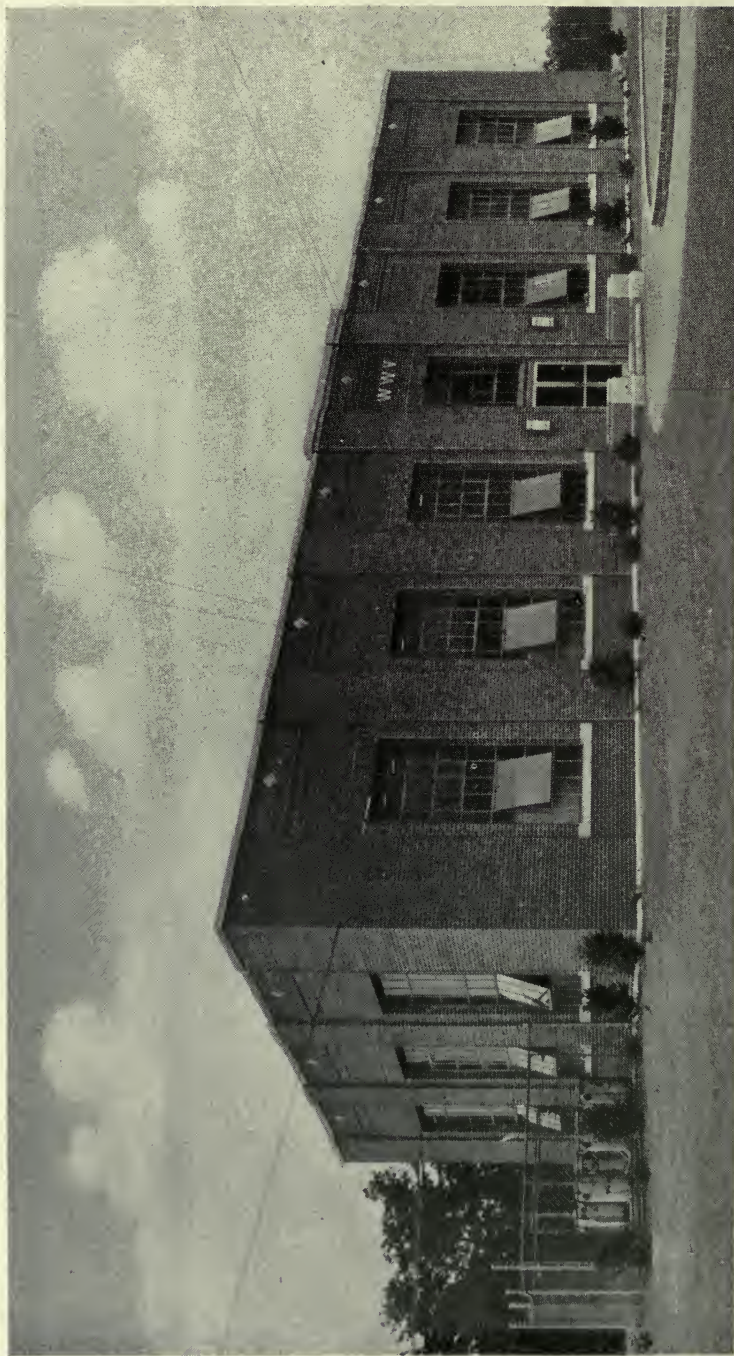


Fig. 32—WWV—The Standard Frequency Broadcasting Station of the U. S. Bureau of Standards.

at Washington, where time determinations of great accuracy are made by means of a Photographic Zenith Tube and a set of quartz clocks. A continuous precise check is maintained between these organizations by radio communication so that the Naval Observatory time signals sent out from NSS at Annapolis and other Navy stations, and from WWV the Bureau of Standards radio transmitting station at Beltsville, Md., as well as all the carrier frequencies from Beltsville, are very accurately determined and maintained in agreement throughout.

The time studies of the U. S. Naval Observatory up to 1937 are described in two important articles by J. F. Hellweg, then Superintendent of the Observatory. The first of these¹¹⁵ in 1932 describes the state of the art just before the quartz clock entered the scene, and the second⁸⁸ in 1937, already referred to, tells of some of the first improvements brought about by its use including the elegant method for making direct photographic time-star checks of the crystal clock rate by means of the Photographic Zenith Tube. Many of the advances involving the use of quartz clocks at the Naval Observatory have not as yet been published.

The British Post Office and the National Physical Laboratory with laboratories at Dollis Hill and Teddington respectively, in cooperation with the Royal Observatory at Greenwich, have done much the same sort of thing in England in relation to time and frequency measurements and broadcast services as has just been described. Considering the number of crystal units among these organizations and the precise nature of the intercomparisons maintained between them, this is probably the most extensive and elaborate quartz clock system in the world. In connection with Greenwich Observatory alone, the complete installation includes eighteen or more such clocks used in deriving the best possible mean rate from stellar observations at Greenwich and from studies of other time observatories throughout the world.

An outline description of the quartz clocks of Greenwich Observatory, and of their function there, has been discussed by Humphry M. Smith in *Electrical Times*¹¹⁶ (London) in March 1946. These clocks employ for the most part the GT cut crystal, first described by W. P. Mason, the bridge stabilized oscillator circuit developed by L. A. Meacham, and the regenerative modulator type of frequency dividers similar to those first developed by R. L. Miller.

The accuracy of the quartz clocks exceeds that of the best pendulum clocks with the result that quartz clocks are now used exclusively in the most precise measurements of time. Some of the considerations¹¹⁷ leading up to the adoption of quartz clocks at Greenwich were discussed in 1937 by H. Spencer Jones, Astronomer Royal. Since then, reports have appeared from time to time by the Astronomer Royal^{89, 118} and others¹¹⁹ concerning the

adoption and use of quartz clocks there. Some interesting sidelights on this "Precision Timekeeping Revolution" were written by F. Hope-Jones in two articles¹²⁰ for the *Horological Journal* during the same year. The quartz clock itself, as developed by the British Post Office for Greenwich Observatory, was described¹²¹ in some detail by C. F. Booth in the *P.O.E.E. Journal* for July 1946. A more general treatment involving some of the same apparatus was presented¹²² by C. F. Booth and F. J. M. Laver in the *I. E. E. Journal* of the same month.



Fig. 33—Crystal chronometer for geophysical studies, consisting of 100 KC. GT-cut crystal, bridge oscillator, and frequency converters to derive precision 500-Cycle output to operate timing devices.

An outstanding example of the versatility of the quartz clock has been its application to the measurement of gravity at sea. Knowing of its stable properties and its independence of gravity, Dr. Maurice Ewing in December 1935, asked the Bell Telephone Laboratories whether a portable quartz clock could be made available for use during a proposed gravity measuring expedition by submarine in the West Indies. Since this was in line with experimental work already in progress at the time, the first portable "crystal chronometer", shown in Fig. 33, was assembled for this occasion, and was taken by Ewing and his colleagues in the U. S. Submarine *Barracuda* on the trip^{80, 81} which began at Coco Solo on November 30, 1936. This was the first application of the GT crystal and the bridge stabilized oscillator in

portable equipment. This original crystal chronometer has been on several gravity-measuring expeditions and is still in active service, having been used again under Dr. Ewing's direction during the summer of 1947.

Gravity determinations at sea are made by measuring the rate of a special triple pendulum that was invented by F. A. Vening Meinesz especially for use in unsteady environments¹²³. Previously, the standard of rate had been the usual ship's chronometers, but Ewing found the crystal chronometer to be an improvement for his purposes, saying in part: "This chronometer is not thermostatted, and temperatures in a submarine change greatly during a dive. No elaborate control over battery voltages was used. The cruise started in the tropics and ended in Philadelphia in mid-winter. It is highly significant that the interval between NAA-time and the chronometer-time never exceeded 0.6 second during the six-week's cruise and that the variation in this interval is very regular. The crystal chronometer has reduced errors in gravity-measurements at sea, due to the rate of the chronometer, to the point where they are negligible."

Some years previous to the construction of the crystal chronometer, a self-contained quartz clock was made to illustrate the possibility of a compact assembly, but it was not sufficiently portable for the submarine expedition. This earlier clock was regulated by a quartz sphere such as used by 'crystal gazers'. The frequency of the sphere was not adjusted, but its natural frequency, which happened to be 33212, was adopted to operate a mean-time dial by the choice of a suitable gear train. Since that time much more compact assemblies have been built using more suitable crystals for control.

The stable properties of the quartz clock have been useful in a number of cases requiring precise synchronization. Perhaps the most noteworthy among these is the application to Long Range Navigation known as LORAN. In this application, pairs of transmitting stations, usually on shore and separated by accurately known distances, send out distinctive signals in synchronism. The time interval between these signals, as received by a ship, identifies the locus of all the points corresponding to that time interval. The set of curves corresponding to all feasible time intervals defines one of the coordinates in a two-coordinate system. The other coordinate is provided in identical manner by another pair of shore transmitters (which may have one station in common with the first pair). The resulting coordinate system consists of two families of intersecting hyperbolas. From the geometry of these curves, and the constants of the signals, the complete figure bounded by the ship and the transmitters can be determined readily.

The need for stability is evident from the fact that the relation between time error and location error is roughly 5 microseconds per mile. In some cases, location within a mile is highly desirable even at considerable dis-

tances. Sometimes the two shore stations, operating as quartz clock time transmitters, must operate for hours without intersynchronization, which calls for very great constancy of rate. One microsecond per hour corresponds to one part in 3.6×10^9 .

The precise synchronization of mechanical parts in remotely situated stations can be accomplished readily. For a number of years, the 5-band privacy system of the transatlantic radio telephone service has been thus synchronized, the apparatus at the American terminal being controlled by the Bell System Frequency Standard while that at the English terminal is controlled independently by similar equipment in the British Post Office. The accuracy requirement for this particular purpose is not very great. However, it has been found possible to maintain two or more rotating shafts at remote and independent stations so precisely controlled by independent quartz oscillators that they never depart, during hours of operation, by more than one fifth of one degree of arc.

A major project in which the quartz clock is destined to take an important part is that of making world-wide land and water surveys in order to locate more accurately boundaries and other features of the earth's surface. There would be applications to sea and air navigation and it would be of great value to geophysicists in studying the figure of, and changes in, the earth's surface. By the combination of a widely dispersed set of Photographic Zenith Tubes associated with quartz clocks and time signal means for communication, and with the powerful ranging techniques growing out of LORAN and RADAR, it should be possible to obtain a new order of accuracy in long distance surveying.

The new order of accuracy of time measurement has made it possible for the first time to study directly the variations in longitude caused by the irregular wandering of the poles. These are small effects and heretofore could only be determined by inference from observations of apparent latitude variations at remote stations. With the added new techniques it should be possible to learn a great deal about these and other phenomena related to real or apparent variations in longitude.

Two other possible applications, involving the precise control of angular movement so readily obtainable with synchronous motors operated from quartz crystal controlled alternating current, are of considerable interest. The first is that of operating the right ascension control of a telescope directly from the amplified output of a crystal-controlled low frequency. Vacuum tube amplifiers and synchronous motors are commercially available with which this could be accomplished by suitable gearing. In addition, of course, it would be necessary to include auxiliary controls to allow for atmospheric and other transient effects, and for obtaining rates of motion

other than sidereal. For small and slowly changing effects this could be taken care of very simply by means of electrical circuits now well known for adding or subtracting small changes in the control frequency.

The other application refers to a suggestion made by the author a few years ago¹²⁴ for the measurement of gravity, and changes in gravity, by comparison of the forces Mg and $M\omega^2 R$. The proposal was based on the idea that ω can be measured or produced with an accuracy two or more orders greater than required, and that the problem reduces to that of balancing two forces and of measuring a linear displacement. The physical set-up would be some form of conical pendulum driven at constant angular velocity about the vertical axis under control of a crystal. Some such arrangements are shown in the reference.

FUTURE POSSIBILITIES

It is part of the nature of a scientist to extrapolate ahead of any current development and to wonder what lies beyond. That feeling is certainly justified in the field of time measurement, for the major advances have taken place in so short a period and so recently, as compared with the thousands of years during which Man has been time-conscious in some degree, that it is reasonable to expect continued advancement for many years to come. Such advancement may come as improvements and refinements in existing techniques, or radically new methods may be developed with inherently more stable potentialities.

Accuracy of Rate

In the first place, it is not reasonable to suppose that the final accuracy that can be attained with the quartz crystal clock has been reached; in view of the rapid current progress indicated in the chart of Fig. 1, it is much too soon to assume this, and there is considerable evidence that improvements could be made by making fuller use of some of the stable properties of quartz crystal and of refinements in the mounting and sustaining circuits. The quartz oscillator assemblies in most general use at the present time embody some compromises which it would not be necessary to make if an all-out effort were being made to construct a few clocks having the highest attainable stability under the most favorable conditions of operation.

The first of these concerns the shape and size of the resonator itself and is related to the frequency of oscillation. From the standpoint of stability of operation, the actual frequency that is used in the oscillator is of little concern because it is now a very simple matter to obtain low frequencies, suitable for the operation of mechanisms, starting with any frequency that can be controlled by a crystal resonator. The choice of 100,000 cycles for the first zero-coefficient resonator was made because, as a standard of frequency,

that value was a good median for the range of frequencies then used in electrical communication. For use in a clock any other frequency would answer just as well, so the inherent stability of the resonator should be given first consideration.

One of the inhibitions imposed on the design of quartz resonators has grown out of the dwindling available supply of large pieces of perfect crystal quartz. Where large quantity production is involved this is an important consideration, but for the small numbers required in a few observatories and national laboratories it should not be a limiting factor.

Except for whatever added difficulties might be entailed in the mounting, it seems reasonable that a large resonator should be more stable than a very small one. The most fundamental reason for this is the proportionate change in effective size that would result from the transfer of any surface material including even the quartz itself.

Every substance is supposed to have some vapor pressure although in some cases it is very minute. However, we are concerned with very minute effects, and it is worthwhile to consider what would happen if there were any evaporation or condensation of material. The possibility of this being an important effect is evident when we realize that the removal of a single layer of molecules from the end of a resonator one centimeter long would increase its frequency by about five parts in a hundred million. The effect on frequency would vary about inversely as the effective length, which favors a large crystal. Such a transfer of material could be inhibited to some extent by operating at a low temperature and by seeking equilibrium between the quartz material of the resonator and other quartz material within the same envelope. Of course, other materials than quartz may be involved in similar surface phenomena and should be thoroughly studied and controlled. This has a strong bearing, of course, on the use of conductive materials deposited on a resonator for the purpose of electrical coupling to it.

The slightest trace of surface contamination has a deleterious effect on the damping coefficient. Professor K. S. Van Dyke in 1935 made a series of measurements on resonators of uniform shape and size but constructed with a considerable range of surface treatments⁴⁵. In the construction of different resonators used in these tests he used different grades of abrasive and various amounts of etching with hydrofluoric acid. In these experiments he operated them under varying degrees of refinement with regard to contamination of the surfaces and found that the highest Q was obtainable only after the utmost care was exercised in keeping the surfaces free from foreign material. The effect is so striking, in fact, that it leads one to wonder whether there is *any* actual elastic hysteresis in the material of quartz crystal, or whether the minute energy losses observed are entirely

surface and coupled effects. Since, for a given shape, the volume increases with linear dimension in greater proportion than the surface area, it can be inferred that surface phenomena would affect a large resonator less than a smaller one.

This is also a reason for employing a stubby shape, in order that the volume of crystal may bear as large a ratio as possible to its surface area. From this standpoint alone a sphere would be ideal but for other reasons, chiefly concerned with the temperature coefficient, it would be unsuitable. It is probable that a polished prolate spheroid, properly oriented with respect to the crystal axes, would satisfy both conditions. Such a resonator could be supported by a pair of wires, serving also as electrical leads from metal-plated electrodes, using techniques already well established.

Crystal resonators as now used in many of the most stable oscillators have been constructed to withstand severe mechanical shock while in operation. It is likely that a slight improvement in frequency stability might be obtained by relaxing a little on the mechanical stability of the present support. Where the greatest accuracy of rate is desired, such as in national standards laboratories and in astronomical observatories, it should be possible to provide suitable mountings for crystal resonators having more delicate supports than those required in mobile equipment. The GT crystal illustrated in Fig. 21 is mounted on eight supporting wires for applications requiring great mechanical stability, and at the same time remains one of the most stable frequency controlling resonators ever produced. It would be reasonable to expect a little improvement in frequency stability at the expense of some mechanical stability if four supports were used instead of eight.

There is a good possibility also that some improvement could be obtained by reducing the electrical coupling to the crystal. At present, the plates are usually provided with plated metal electrodes which cover the entire large surface areas. Some increased stability in frequency might be expected by the use of relatively smaller electrodes covering only the central part of the resonator where the amplitude of vibration is small. At least two advantages might be expected from such a modification. One is that the loading effect is least near the node for vibration, another is that any looseness of material, or elastic hysteresis, would be least troublesome where the motion is least. Of course, it is chiefly the *variations* in such effects that concern us. One would expect, however, that if such effects exist at all they might be minimized by the use of smaller electrodes.

These particular effects may be eliminated completely, of course, by the use of isolated electrodes spaced from the crystal—but at the expense of other possible variations related to changes in electrode spacing. There is

considerable promise in such means, the end result depending upon how precisely the resonator may be held in a fixed position by means that will not change its resonance characteristics. Such means have, in fact, been used successfully in a number of German quartz clocks such as at the Physikalisch-Technische Reichsanstalt¹⁰⁸, and with the Dye ring resonator developed by D. W. Dye and L. Essen at the National Physical Laboratory^{125, 126}, England.

For any given resonator and circuit a careful study would probably reveal an optimum amplitude of oscillation that would yield a maximum stability against residual uncontrollable variables. With the GT crystal, as used currently, the maximum amplitude of motion is about 0.00006 mm. It would be possible to limit the motion to a tenth or a hundredth of this value if it should be found desirable.

Further studies of the factors contributing to aging of the quartz material also should produce valuable improvements. Since resonators, which appear to be alike in all other respects, often age at greatly different rates, some being very small or substantially zero, it would seem that some reason should be discoverable for such variations and some effective control established.

There are other relatively massive shapes that should be investigated further such as the ring crystal, mentioned earlier in this paper, and as developed and studied by Dye and Essen^{125, 126}. The ring may be excited in various modes of vibration some of which are more favorable than others from the standpoint of mounting. By choice of orientation relative to the crystal axes, and of dimensions, certain of these can be designed to have zero temperature coefficients in a restricted temperature region.

Another shape that holds great promise because of its convenience of mounting, along with the other desirable properties, is the rectangular rod vibrating longitudinally in its second or higher overtone such as first described by Scheibe and Adelsberger¹⁰⁸. Still another possible massive shape is a much thicker version of the GT crystal which would combine the very favorable temperature-frequency characteristic with that of reducing the ratio of surface area to volume.

In seeking the highest possible accuracy a precise temperature control is essential in all cases, even with the GT type of resonator with its wide region of low-temperature coefficient. The reason for this is that the frequency of oscillation depends not only on the mean temperature of the resonator but also upon the temperature gradient throughout its volume. Thus, even if a resonator has the same frequency exactly at different mean temperatures, its frequency will vary a little while the temperature is varying from one value to another. The effect of this can be reduced by enclosing the crystal unit in an envelope with thermal lagging so that such *variations* as do exist at the temperature control layer are prevented from reaching the crystal.

This is no longer a serious problem for there are various electronic means such as described by C. F. Booth and E. J. C. Dixon¹²⁷ for continuous temperature control, by means of which the variations may be kept very small, and very effective thermal lagging methods¹²⁸ are well known.

The bridge method for temperature control has been applied in many forms. One of the simplest and most effective procedures has been to utilize a bridge-stabilized oscillator of the type developed by L. A. Meacham for *frequency* control, and to use it instead for *temperature* control. For this purpose, all four arms of the bridge are noninductive resistances wound as heaters on the oven to be controlled. In the feedback circuit of the oscillator, a rough frequency control is included simply for the purpose of setting up an oscillation in the circuit which includes the bridge. The conjugate pairs of bridge arms are made of resistance wire with different temperature coefficients and so proportioned that the bridge balances at the desired temperature. The *amplitude* at which this bridge oscillator oscillates depends upon the temperature departure from the balance value. Since the alternating current output of the oscillator flows in the bridge arms, the amount of heating is proportional to the temperature error, and hence the control is automatic.

Continuity of Operation

An astronomical clock, in addition to having as nearly constant a rate as can be attained, should also be able to operate over long periods of time without change or interruption. The reason for this is that many of the phenomena that are of interest in time measurement occur in continuous succession and the greatest amount of information can be obtained only by the use of clocks with which measurements can be made in unbroken sequence. Quartz clocks that have been used for astronomical purposes to date have not had a very commendable record in this respect and already a good deal has been said in the clock literature about this aspect—as though it were an inherent property of the quartz clock.

However, it is only a matter of simple engineering, making use of techniques and apparatus already well known and available, to design a quartz clock which should operate continuously for many years. A chain is only as strong as its weakest link—and the clock comprises a chain of apparatus parts every link of which must function perfectly and continuously. This chain consists of (1) the crystal-controlled oscillator, (2) a frequency demultiplier to obtain a low frequency to operate a motor, (3) a power amplifier to obtain sufficient current to drive the motor and (4) the motor itself, associated with any of a wide assortment of time signal-producing or measuring equipment. In addition to the links in this chain, a power supply must be maintained, and the temperature of the crystal must be controlled, both continuously.

The crystal itself is no problem as far as continuity of operation is concerned. Its motion is so very small there is no likelihood at all of failure on that account. Mountings are very stable and in all likelihood will be improved. The oscillator circuit, the frequency demultiplier, the power amplifier and the temperature-control circuit are all vacuum-tube devices and deserve special consideration. In all of these circuits, vacuum tubes have been used in some installations which do not have a very long life, some even becoming defective within a year of operation. On the other hand, there are tubes which have been developed for use in continuous telephone circuits where failures would be troublesome and costly. Some of these tubes in current production have an expected life of more than ten years. There is good reason to believe that a quartz clock installation equipped with such vacuum tubes throughout, and engineered so as to make effective use of their special properties, would operate continuously for ten years or more.

The remaining "link" in the chain is the synchronous mechanism operated from the crystal-controlled circuits and used for totalizing continuously the oscillations of the crystal and for producing suitable time signals at specified intervals of time thus measured off in terms of the crystal rate. This mechanism usually consists of a small synchronous (phonic wheel) motor operated from a submultiple of the crystal frequency and geared to commutators or cams or other means for producing the electrical signals used in making time measurements. Many of the troubles in quartz clock installations have occurred in this 'link'. There is every reason to believe, however, that suitable synchronous motors geared to cam-controlled electrical contacts can be built that will operate continuously through many years. To insure long operation it would be desirable to employ motors with low rotation speed in order to reduce bearing wear. With the present knowledge of bearing materials and lubricants, it should be a simple matter to design such a motor that would operate without failure for ten years or more.

A relatively trouble-free electrical time signal producer, suitable for operating under the control of a quartz oscillator, with frequency demultipliers to 100 cycles, could be constructed as indicated schematically in Fig. 34. This is not intended to be an actual design, but is intended to indicate how an apparatus could be designed that would circumvent some of the troubles now experienced which prevent long continuous operation.

The basic apparatus consists of a crystal oscillator, presumably 100,000 cycles, with a frequency divider to obtain controlled 100-cycle current to drive the 100-pole phonic wheel motor at one revolution per second. Obviously, other crystal frequencies and step-down ratios could be used, the

important thing being to obtain a rotation speed of 1 rps. This is a very low speed for a phonic wheel motor but has the obvious advantage of great simplicity since it permits of controlling seconds devices without the use of gearing. Only one shaft is involved and the bearing problem is reduced to the simplest possible terms. A hardened steel cam, integrally mounted with the phonic wheel rotor, is used to operate a single electrical contact, so connected into the circuits controlled by it that the *instant of break* is the sole time-determining operation. A *break* signal is preferable to a *make* signal chiefly because it is easier to avoid irregular effects, such as result from contact chatter, when a circuit is being opened than when it is being closed. If a pallet of sapphire or ruby is used for the mechanical contact on the cam,

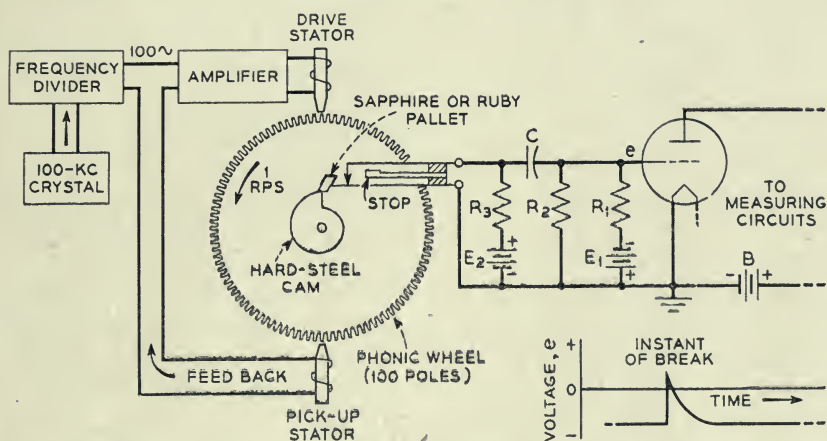


Fig. 34—Suggestion of elements for a quartz clock for long time continuous operation.

and if small currents are used through the contacts, made preferably of platinum-iridium or similar alloy, it would be reasonable to expect trouble-free performance through several hundred million operations.

Ordinarily, the "hunting" of a phonic wheel motor operating on a frequency as low as 100 cycles would cause time errors too large to neglect in a device such as just described. However, by the use of feedback in the motor amplifier circuit, such as indicated schematically in Fig. 34, the effective hunting can be reduced to the point where the time errors caused by it would become negligible for most purposes.

Various circuits could be suggested for making use of the break signal for timing purposes, the one shown in Fig. 34 being typical and suitable for various methods of precise measurement and control. It is capable of providing an electrical impulse with a steep wave front and of adjustable duration. The grid of the vacuum tube is normally biased to cutoff by the

negative voltage, $\frac{E_1 R_2}{R_1 + R_2}$. While the contact is closed, the battery E_2 , with resistance R_3 in series, is short-circuited. But at the instant of opening the contact, current flows momentarily in the circuit including E_2 , R_3 , C and R_2 . By making E_2 positive, and equal to or larger than E_1 numerically, the plate circuit of the tube becomes conducting for a short interval, the duration of which is determined by the time-constant of the condenser circuit, each time the contact is opened. At all other times, the plate circuit is nonconducting. The sharply defined electrical signal thus produced in the plate circuit can be used by well-known means for direct time comparison with signals from other sources.

Making use of the *duration* of the impulse thus produced, it is possible to use it as a selecting means to isolate a single more precise signal from a continuous chain. For example, the 100-cycle wave controlled by the crystal can be modified by a simple vacuum tube circuit to consist of a continuous sequence of very sharply defined impulses. By using the pulse circuit just described as a bias control on an amplifier, it would be readily possible to select one out of every hundred of these impulses and thus provide an extremely precise seconds signal, the accuracy of which is determined wholly by electronic means.

It would be readily possible to vary the time relation of the seconds signal while in operation, by the use of electrical phase shifters in the driving circuits, or by rotating the stator of the phonic wheel motor, but for long continuous operation it would be desirable to keep the number of apparatus parts comprising the clock at a minimum.

It is not necessary, of course, to employ a complete frequency divider and phonic wheel apparatus for each quartz crystal oscillator. As mentioned previously, the relative time rates of quartz oscillators can be measured with very high precision and be very simple means through a direct comparison of the high frequencies.

Other Means for Precise Rate Control

In addition to making improvements on the quartz crystal resonator, and on methods for sustaining it in vibration, there are two other avenues of investigation which may yield comparable results, with possibly some additional advantages. Not much can be said about them at this time except to point out their possibilities because no appreciable work has been done so far to explore their merits as timekeepers.

The first is in the field of very low temperatures where some quite remarkable properties are obtained. Chief of these for our purpose is the supraconductivity of some metals, and the constancy of shape of most materials, at temperatures in the neighborhood of absolute zero. It seems

reasonable to suppose that an electrically-resonant circuit maintained at a temperature in this region could be made to have a very high Q , and very stable dimensions, and so have the chief desirable properties for rate control that obtain in a quartz resonator. Resonant cavities used at high frequencies have many of the properties of other electrical resonant circuits, and in particular their energy dissipation for electric oscillations can be very substantially reduced when cooled to superconducting temperatures. In some experiments made recently at Massachusetts Institute of Technology¹²⁹ it has been shown that a cavity resonator made of lead, which for 3-cm. waves has a Q of about 2,000 at room temperatures, is so much improved at a temperature of 4 degrees absolute that the Q approaches a million. Such a resonator could be used as the stabilizing element in an oscillator and hence in a clock. The relative stability over long periods could, of course, be determined only by experiment.

Maintenance of the required low temperature would add considerably to the complexity of such a system, but if the advantages were such as to produce a new order of stability, and particularly if it should make possible a clock system with small or zero aging, it certainly should be justified for future time measurement studies.

The other avenue of approach is through the application of certain resonance phenomena in atoms and molecules that do not depend upon aggregates of matter as is the case with all mechanical systems used heretofore in time measuring means. The extreme fineness of structure and the constancy of atomic and molecular resonance phenomena have long been recognized through studies of line spectra, and in the field of spectroscopy these properties have been used as standards of wavelength ever since the early studies of Joseph von Fraunhofer, reported in 1815.¹³⁰ Wavelength, λ ,

and frequency, f , are associated by the simple relation $f = \frac{c}{\lambda}$ where c is equal to the velocity of light. For visible radiations f turns out to be extremely large, for the red light, 6500Å, it is 462 million million vibrations per second. So far, such high frequencies have not been observable or measurable directly but can only be deduced from wavelength measurements as just stated—which inevitably involve the use of man-made standards of length and the combined errors of two quite different sorts of physical measurements.

It has long been the dream of physicists to find some way to tie in directly with the natural frequencies of atoms and molecules and to derive from them a direct measure of rate, and, of course, of time interval. It has been thought, for example, that the red radiation from cadmium vapor, whose wavelength was measured by C. Fabry and A. Perot in terms of the standard meter as accurately as that standard could be defined, would also make a

good standard for time measurements. A step in the right direction was made later by A. A. Michelson whose precise determination of the wavelength of this radiation made possible the redefinition of the International Meter as a definite number of such wavelengths, measured in vacuo. From this definition, it is now possible to duplicate the primary standard of length with great accuracy, and to check such secular changes as may occur in the original standard, the distance between two marks on a metal bar. The constancy of the standard, as defined by Michelson, depends upon properties of primary particles of matter, and upon properties of space, which, as far as human beings are concerned directly, appear to be quite independent of time or location. A similar definition of rate, or time interval, is very desirable.

A ray of hope came out of the important work of Nichols and Tear¹³¹ who proved that electric waves which could be produced electrically were of the same stuff as radiation from hot bodies. They were able to detect radiation of either sort by the same receiving device and showed that they both had the same properties of refraction, polarization, etc. Later, Cleeton and Williams¹³² were able to produce *continuous* electric waves at very high frequencies—corresponding to about 1 cm. wavelength—and to show that they also had the important properties of light waves. Now the range has been extended somewhat more and there are reports¹³³ of experimental generators that can produce continuous waves of a few millimeters wavelength. This is an active development and, of course, the end is not in sight. From continuous waves of any frequency it is believed possible by general techniques now well known to control lower frequencies, and from them eventually all sorts of time measuring and indicating devices as previously described.

Within the last few years, the missing link has been discovered which, with suitable instrumentation, may make it possible to construct a clock controlled by atomic- or molecular-resonance phenomena. There are a great number of resonance phenomena associated with the molecules in a gas, or in molecular beams, which are responsive to electric waves that can be produced continuously by modern vacuum tube means. In some cases, the sharpness of resonance is such that changes of frequency of one part in 10^8 or less can be detected, leading to the idea that such resonance phenomena may be utilized in some way to *control* the frequency of a suitable oscillator and hence, through frequency conversion circuits, to control frequencies low enough to operate clocks and other mechanisms. Some of the resonance phenomena in point are in the one-centimeter region, a field that is rapidly being exploited in radar and communication applications. It is to be expected, therefore, that techniques for dealing with such high frequencies will be developed in the near future thus facilitating a study of this new

approach to timekeeping. The idea of utilizing such resonance phenomena for the measurement of time was suggested in January, 1945 by Professor I. I. Rabi of Columbia University at an address before the American Physical Society and the American Association of Physics Teachers.

These resonance phenomena, involving the interaction of microwave electromagnetic radiation with atoms or molecules of matter, have been discovered only quite recently and it is likely that a great deal more will be learned about them in the next few years. The results already obtained are very promising and investigations already under way may well lead to the means for creating an entirely new type of standard of time interval and rate—both of prime importance in Physics.

The studies of greatest significance for such purposes now in progress fall in two main branches involving quite different techniques. The actual means for regulating a clock would be quite different in the two methods, but would be possible in either. With what is known up to the present time, however, the construction of such a clock would be a considerable undertaking, especially to make one that would operate over long periods. The two chief phenomena involving atomic or molecular resonances are: (1) the absorption of high-frequency energy in certain materials, particularly in gases, exhibiting ultra-fine absorption spectra; and (2) the deflection of beams of atoms or molecules under special conditions of magnetic and electric fields. The earliest reported work on the absorption of microwaves in gases was done by C. E. Cleeton and N. H. Williams¹³⁴ in 1934. With the development of improved high-frequency generators and measuring techniques the work has been extended considerably during the last few years by C. H. Townes¹³⁵, W. E. Good¹³⁶ and others. It is believed that with modifications of methods, such as used by them, it would be possible to control the frequency of the short-wave generators such as used in making these studies; and, if this can be done, the adaptation for use in time-measuring devices would follow naturally as in the case of any other stable oscillator.

The general method using molecular beams has been a gradual development over some years, but the first published suggestion of the applications which relates closely to this work was made in 1938 when I. I. Rabi, J. R. Zacharias, S. Millman and P. Kusch first used the beam deflection method for measuring nuclear magnetic moments.¹³⁷ Two articles^{138, 139} in *Reviews of Modern Physics* in July 1946 give a good description of the molecular beam method and the results of some studies of fine structure resonance phenomena. The resonance curve shown in Fig. 35 obtained recently by P. Kusch and H. Taub of Columbia University, and hitherto unpublished, illustrates the resolution obtainable by molecular beam methods. According to theory, the actual *width* of the resonance should be substantially inde-

pendent of the applied frequency and they expect to be able, when employing frequencies corresponding to centimeter waves, to obtain a hundred or more times this resolution. If this should be realized, it suggests the possibility of a clock with an accuracy of better than one part in 10^8 .

Perhaps the greatest advantage that might be expected from such a method lies in the possible long-time stability or freedom from aging. Every existing means for timekeeping involves in some manner the motion of large aggregates of matter which, when they rearrange themselves in any way, vary their rates of rotation, or of oscillation, as the case may be, in ways

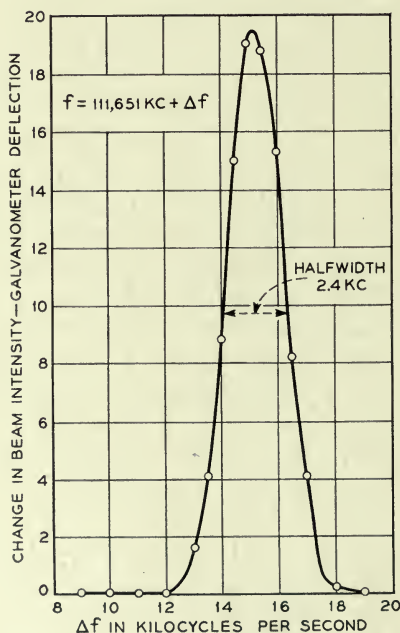


Fig. 35—Typical resonance curve for a line in the radio frequency spectrum of atomic K^{39} observed by the method of molecular beams. Experimental data supplied by P. Kusch and H. Taub, Columbia University Physics Department.

that are not wholly predictable. It may well develop that a method based on the behavior of single particles of matter will be ageless and, with proper instrumentation, that it will permit of setting up an *absolute* standard of rate and time interval. The actual value of this rate would be indeterminate by a small amount depending on the sharpness of resonance and the precision of control that could be effected from it, in addition to any uncontrollable effects of the actual resonance frequencies such as result from temperature, pressure, and electromagnetic and gravitational force fields. In the case of some of the resonance phenomena all the latter effects are believed to be

vanishingly small. In any case, one would not expect to experience a progressive change in rate as in the case of the rotation of the earth which now is the measure and definition of astronomical time. On the average the earth is said to be slowing down at the rate of a thousandth of a second per day per century¹⁴⁰ and, according to the astronomers⁶⁹, the day will continue to lengthen until finally, at some time in the distant future, the earth will always face one side toward the moon and the length of the day will become about 47 times as long as it is at the present time.

Meanwhile, if an absolute standard could be established, such as now appears feasible through atomic- or molecular-resonance phenomena, it would be possible to record these changes through the centuries and to establish a relatively stable "second" that could be used for all time in physical measurements in place of the elastic second of the cgs system which, as now defined, must stretch with the inevitable variations in the mean solar day.

Whether or not such an "absolute" clock becomes a reality at some time in the future, the quartz crystal clock, because of its accuracy, compactness, great convenience and versatility is likely to continue to be a most useful instrument in all precision measurements of time.

REFERENCES

1. The Beginnings of Time-Measurement and the Origins of Our Calendar. James Arthur Foundation Lecture), James Henry Breasted. Published in Book, Time And Its Mysteries, 1936. New York University Press.
2. Encyclopedia Britannica—14th Ed.—"Christian Huygens".
3. Horologium Oscillatorium, Christian Huygens, 1673.
4. Electrical Timekeeping, F. Hope-Jones, N. A. G. Press, London, 1940.
5. Improvements in the Application of Moving Power to Clocks and Timepieces, John Barwise and Alexander Bain, British Patent No. 8783, Filed Jan. 11, 1841, Issued July 10, 1841.
6. The first electric clock, W. A. Marrison, *Proceedings of the Engineering Society*, Queens University, Kingston, Canada, v. 29, pp. 15-20, 1940.
7. Electric Clocks, F. Hope-Jones—Book, N. A. G. Press, London, 1931.
8. Memoire sur l'étude optique des mouvements vibratoires, Jules Lissajous, *Comptes Rendus*, v. 44, p. 727, April 6, 1857.
9. On the use of the dynamic multiplier with a new accompanying apparatus, C. G. Page, *American Journal of Science*. 1st Series, 32, 1837, p. 354, dated at Salem, Mass. April 24, 1837.
10. Observations on induced electric currents with a description of a magnetic contact breaker, Golding Bird, London, Edinburgh and Dublin, Phil. Mag. Series 3, No. 12, 1838, p. 18. Addressed from Wilmington Square, Nov. 2, 1837.
11. On a new magnetic electric machine, *Annals of Electricity, Magnetism and Chemistry and Guardian of Experimental Science*, vol. III, pp. 66-70, 1839. Translation from German. Describes apparatus of Dr. Neeff of Frankfurt exhibited at Friburg meeting of Philosophers, September, 1838.
12. Science of Musical Sounds, Dayton Clarence Miller. Book, Macmillan, N. Y. 1916. p. 29.
13. Quelques Experiences d'Acoustique, Rudolph König, Paris, 1882. p. 172.
14. On the characteristics of electrically operated tuning forks, H. M. Dadourian, *Physical Review*, v. 13, pp. 337-359, May, 1919.
15. Isochronous and Synchronous Movements for Telegraph and Other Lines. Patent No. 203423, Poul la Cour. Filed April 9, 1878.

- 15a. Roue phonique pour la régularisation du synchronisme des mouvements. Note by P. la Cour in *Comptes Rendus*, v. 87, pp. 499-500, September 25, 1878.
16. Report of Physical Society Meeting of March 30, 1878. Refers to Lord Rayleigh's impulse motor. *Nature*, May 23, 1878, p. 111.
17. Nouveaux modes d'entretien des diapasons, A. and V. Guillet. *Comptes Rendus*, v. 130, pp. 1002-1004, April 9, 1900.
18. The Emission of Electricity from Hot Bodies, O. W. Richardson. Book. Longman, Green and Co., London, 1916.
19. On electric discharge between electrodes at different temperature in air and high vacuo, J. A. Fleming, *Royal Society of London Proc.*, v. 47, p. 122, 1890.
20. The Audion—A New Receiver for Wireless Telegraphy, Lee De Forest, *A. I. E. E. Transactions*, v. 25, pp. 735-779, 1906. *Electrician*, v. 58, pp. 216-218, 1906.
21. Improvements in instruments for detecting and measuring alternating currents, J. A. Fleming. *British Patent No. 24850*, 1904.
22. Device for amplifying feeble electrical currents, Lee De Forest. *U. S. Patent No. 841387*, issued January, 1907.
23. Einrichtung zur Erzeugung elektrischer Schwingungen, Siegmund Strauss. *Austrian Patent No. 71340*, filed Dec. 1912, issued June, 1915.
24. Improvements in receivers for use in wireless telegraphy and telephony, Marconi Wireless Telegraph Company, Ltd. and Charles Samuel Franklin. *British Patent No. 13636*, filed June, 1913, accepted June, 1914.
25. Wireless Telegraph and Telephone System, Lee De Forest. *U. S. Patent No. 1,507,016*, filed September, 1915, issued September, 1924, and *U. S. Patent No. 1,507,017*, filed March, 1914, issued September, 1924.
26. The use of the triode valve in maintaining the vibration of a tuning fork. W. H. Eccles, *Phys. Soc. of London Proc.*, v. 31, p. 269, 1919.
27. Sustaining the vibration of a tuning fork by a triode valve, W. H. Eccles and F. W. Jordan, *The Electrician*, v. 82, p. 704, June 20, 1919.
28. Sur L'entretien des oscillations mécaniques au moyen des lampes à trois électrodes, Henri Abraham and Eugene Block, *Comptes Rendus*, v. 168, pp. 1197-1198, June 16, 1919.
29. Electron tube drive for tuning fork, E. A. Eckhardt, J. C. Karcher, and M. Keiser, *Physical Review*, v. 17, pp. 535-536, April, 1921.
30. An electron tube tuning fork drive, E. A. Eckhardt, J. C. Karcher and M. Keiser, *J.O.S.A.—No. 6*, pp. 949-957, November, 1922.
31. The valve maintained tuning fork as a precision time standard, D. W. Dye, *Royal Soc. of London Proc.*, v. 103, pp. 240-260, May, 1923.
32. Frequency measurements in electrical communication, J. W. Horton, N. H. Ricker, W. A. Marrison, *A. I. E. E. Trans.*, v. 42, pp. 730-741, June, 1923.
33. A celestial encounter, Ernest William Brown, *Jl. Franklin Institute*, v. 202, pp. 127-163, August, 1926.
34. The best observed eclipse in history, *Scientific American*, March, 1925, p. 155.
35. Precision determination of frequency, J. W. Horton and W. A. Marrison, *I. R. E. Proc.*, v. 16, pp. 137-154, February, 1928.
36. The valve maintained tuning fork as a primary standard of frequency, D. W. Dye and L. Essen, *Royal Society of London Proc.*, v. 143, pp. 285-306, 1934.
37. The anomaly of nickel steels, Charles Edouard Guillaume, *Physical Soc. of London Proc.*, v. 32, p. 374, April, 1928.
38. La compensation des horloges et des montres; précédés nouveaux fondés sur l'emploi des aciers au nickel, Charles Edouard Guillaume. Booklet—Neuchatel et Genève, Paris, 1920.
39. Dilatibilité du chrome et des alliages nickel-chrome dans un intervalle étendu de temperature, P. Chévenard. *Comptes Rendus*, v. 174, p. 109, January, 1922.
40. Mesure de la dilatation du coefficient thermoelastique et propriétés électriques des alliages dans un grand intervalle de temperature. Résultats de l'étude des ferromagnétiques purs et additionnés le chrome. P. Chévenard. *Bull. Soc. Franc. Phy.* No. 254, pp. 135-138, Dec. 16, 1927.
41. Tuning forks, H. H. Hagland. *U. S. Patent 1,715,324*, filed June, 1925, issued May, 1929.
42. A mechanical oscillator of constant frequency, August Karolous. *U. S. Patent 1,763,853*, filed Nov., 1927, issued June, 1930.
43. Compensated Tuning Fork, Bert Eisenhour. *U. S. Patent 1,880,923*, filed Sept., 1930, issued Oct., 1932.

44. Tuning Fork, S. E. Michaels. *U. S. Patent* 2,247,960, issued July 1, 1941.
45. A determination of some of the properties of the piezoelectric quartz resonator, Karl S. Van Dyke, *Proc. I. R. E.*, v. 23, No. 4, April, 1935.
46. The high Q of quartz resonators, Maynard Waltz and K. S. Van Dyke, *Jl. of Acoustical Soc. of America*, v. 19, No. 4, Part 1, p. 732, July, 1947.
47. The crystal clock, W. A. Marrison, *National Academy of Sciences Proc.*, v. 16, pp. 496-507, July, 1930.
48. Développement par pression, de l'électricité polaire dans les cristaux hemiedres a faces inclinées, Jacques and Pierre Curie, *Comptes Rendus*, v. 91, p. 294, 1880.
49. Déformations électrique du quarts, Jacques and Pierre Curie, *Comptes Rendus*, v. 95, pp. 914-197, 1882.
50. Oeuvres de Pierre Curie, Pierre Curie. Book—Gauthier-Villars, Paris, 1908.
51. The electric network equivalent of a piezoelectric resonator, K. S. Van Dyke, *Phys. Rev.*, v. 25, p. 895, 1925.
52. The piezoelectric effect in the composite rochelle salt crystal, A. McLean Nicolson, *A. I. E. E. Trans.*, v. 38, Part 2, pp. 1467-1485, 1919. *A. I. E. E. Proc.*, v. 38, pp. 1315-1333, 1919.
53. Generating and transmitting electric currents, Alexander M. Nicolson. *U. S. Patent No.* 2,212,845, filed April 10, 1918, issued August 27, 1940.
54. Sondage par le son, P. Langevin. *S. 29, Pub. Spec.* No. 3, 1924.
55. Improvements relating to the emission and reception of submarine waves, P. Langevin. *French Patent No.* 505,903 issued in 1918, also *British Patent No.* 145,691 issued in 1921.
56. The piezoelectric resonator, W. G. Cady. *I. R. E. Proc.*, v. 10, pp. 83-114, April, 1922.
57. Piezoelektrische Resonanzerscheinungen, A. Scheibe, *Zeitschrift für Hochfrequenztechnik*, v. 28, No. 1, 1926.
58. Bibliography on piezoelectricity, W. G. Cady, *I. R. E. Proc.*, v. 16, pp. 521-535, 1928.
59. Quarzoszillatoren, R. Bechmann. *Telefunken Zeitung*, v. 17, pp. 36-45, March, 1936.
60. Piezoelectricity, W. G. Cady. Book—McGraw-Hill, New York, 1946.
61. Piezoelectric crystal resonators and crystal oscillators applied to the precision calibration of wave meters, George W. Pierce, *Amer. Acad. of Arts and Sciences Proc.*, v. 59, No. 4, pp. 81-106, October, 1923.
62. Sur un nouvel emploi des quartz piezoélectriques, G. Siadbei, *Comptes Rendus*, v. 188, p. 1390, May, 1929.
63. Piezoelectric crystal oscillators applied to the precision measurement of the velocity of sound in air and carbon dioxide at high frequencies, George W. Pierce, *Amer. Acad. of Arts and Sciences Proc.*, v. 60, pp. 271-302, October, 1925.
64. Sichtbarmachung von hochfrequenten longitudinalschwingungen piezoelektrischen Kristallstabe, E. Giebe and A. Scheibe, *Zeits. f. physik* v. 33, pp. 335-344, 1925.
65. Leuchtende piezoelektrische resonatoren als hochfrequenznormale, E. Giebe and A. Scheibe, *E. T. Z.*, v. 47, pp. 380-385, April, 1926.
66. An international comparison of frequency by means of a luminous quartz resonator, S. Jimbo, *Proc. I. R. E.*, v. 18, pp. 1930-1934, 1930.
67. An international comparison of radio wavelength standards by means of piezoelectric resonators, W. G. Cady, *Proc. I. R. E.*, v. 12, pp. 805-816, December, 1924.
68. The status of frequency standardization, J. H. Dellinger. *Proc. I. R. E.*, v. 16, No. 5, pp. 579-592, May, 1928.
69. Frequency control system, Warren A. Marrison. *U. S. Patent No.* 1,788,533, filed March, 1927, issued January, 1931.
70. Precision determination of frequency, J. W. Horton and W. A. Marrison, *I. R. E. Proc.*, v. 16, pp. 137-154, February, 1928.
71. Sur la mesure en valeur absolue des périodes des oscillations électriques de haute fréquence, Henri Abraham et Eugene Block, *Comptes Rendus*, v. 168, pp. 1105-1108, 1919.
72. "Universal" frequency standardization from a single frequency standard, J. K. Clapp, *J. O. S. A.*, v. 15, No. 1, pp. 25-47, July, 1927.
73. Frequency demultiplication, Balth Van der Pol and J. Van der Mark, *Nature*, v. 120, pp. 363-364, Sept., 1927.
74. Über Relaxationschwingungen, B. Van der Pol, Jr., *Zeits. f. Hochfrequenztechnik*, v. 29, pp. 114-118, April, 1927.

75. Fractional frequency generators utilizing regenerative modulation, R. L. Miller, *I. R. E. Proc.*, v. 27, No. 7, pp. 446-457, July, 1939.
76. Generation and control of electric waves, Joseph W. Horton. *U. S. Patent No. 1,690,299*, filed March, 1922, issued November, 1928.
77. Piezoelectric crystal, Warren A. Marrison. *U. S. Patent No. 1,899,163*, filed Dec., 1928, issued Feb., 1933. *U. S. Patent No. 1,907,425*, filed Dec., 1928, issued May, 1933. *U. S. Patent No. 1,907,426*, filed Dec., 1928, issued May, 1933. *U. S. Patent No. 1,907,427*, filed Dec., 1928, issued May, 1933.
78. A high precision standard of frequency, W. A. Marrison, *I. R. E. Proc.*, v. 17, pp. 1103-1126, July, 1929. *B. S. T. J.*, v. 8, pp. 493-514, July, 1929.
79. The Crystal Clock, W. A. Marrison, *National Academy of Sciences Proc.*, v. 16, pp. 496-507, July, 1930.
80. Gravity measurements on the U. S. S. *Barracuda*, Maurice Ewing, *Transactions of the American Geophysical Union*, 1937.
81. Gravity at sea by pendulum observations, Albert J. Hoskinson, *American Institute of Mining and Metallurgical Engineers Technical Publication No. 955*, New York Meeting, 1938.
82. Modern developments in precision clocks, Alfred L. Loomis and W. A. Marrison, *A. I. E. E. Trans.*, v. 51, pp. 527-537, June, 1932.
83. The precise measurement of time, A. L. Loomis, *Monthly Notices, R. A. S.*, v. 91, March, 1931.
84. The spark chronograph, W. A. Marrison, *Bell Laboratories Record*, v. 18, No. 2, pp. 54-57, October, 1939.
85. Quartz Crystals for Electrical Circuits, Raymond A. Heising. Book. D. Van Nostrand, Inc. New York, 1946.
86. A new quartz crystal plate, designated the GT, which produces a very constant frequency over a wide temperature range. W. P. Mason. *I. R. E. Proc.*, v. 28, No. 5, pp. 220-223, May, 1940.
87. The bridge stabilized oscillator, L. A. Meacham, *I. R. E. Proc.*, v. 26, No. 10, pp. 1278-1294, October, 1938.
88. Time determination and time broadcast, J. F. Hellweg, *Franklin Institute Jl.* No. 223, pp. 549-563, 1937.
89. The measurement of time, Sir Harold Spencer Jones, *Endeavour*, v. 4, pp. 123-130, October, 1945.
90. Translating Circuits. (A continuous phase shifter), W. A. Marrison. *U. S. Patent No. 1,695,051*, filed May, 1925, issued Dec., 1928.
91. Standard Frequency System, W. A. Marrison. *U. S. Patent No. 2,087,326*, filed Sept., 1934, issued July, 1937.
92. Phase Shifting Apparatus, Larned A. Meacham. *U. S. Patent No. 2,004,613*, filed Aug., 1933, issued June, 1935.
93. A multiple unit steerable antenna for short wave reception, H. T. Friis and C. B. Feldman, *Proc. I. R. E.*, v. 25, pp. 841-915, July, 1937.
94. Clocks showing mean and sidereal time simultaneously, F. Hope-Jones, *Franklin Institute Jl.* No. 223, pp. 95-100, January, 1937.
95. Sur des horloges indiquant simultanément le temps solaire moyen et le temps sidéral, Ernest Esclançon. *Comptes Rendus*, 206, pp. 289-292, January 31, 1938.
96. Electrical Engineers' Handbook, Pender-McIlwain. Chapter on Frequency Measurement by W. A. Marrison. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London.
97. Frequency standard, W. A. Marrison. *U. S. Patent No. 1,935,325*, filed April, 1929, issued Nov., 1933.
98. Frequency measurement, W. A. Marrison. *U. S. Patent No. 1,936,683*, filed Sept., 1931, issued Nov., 1933.
99. High precision frequency comparisons, L. A. Meacham. *The Bridge of Eta Kappa Nu.*, v. 36, pp. 5-8, Feb.-March, 1940; also *Bell System Technical Monograph B-1232*.
100. An instrument for short-period frequency comparisons of great accuracy, H. B. Law, *The Jl. of the Institution of Electrical Engineers*, Vol. 94, Part III, No. 27, pp. 38-41, January, 1947.
101. Electron and Nuclear Counters—Theory and Use, Serge A. Korff. Book. D. Van Nostrand Company, Inc., New York, 1946.
102. Electronic Counters, I. E. Grosdoff, *R. C. A. Review*, v. 7, No. 3, pp. 438-447, Sept., 1946.

103. Method and Means for Indicating Synchronism, W. A. Marrison. *U. S. Patent No.* 1,762,725, filed March, 1928, issued June, 1930.
104. Split second time runs today's world, F. Barrows and Catherine Bell Palmer, *The National Geographic Magazine*, v. 92, No. 3, Sept., 1947. See page 402.
105. Frequency and Time Control Aided by Telephone Company, H. C. Forbes and F. Zauggbaum, *Electrical World*, pp. 117-118, January 20, 1934.
106. Generation of reference frequencies, L. A. Meacham, *Bell Laboratories Record*, v. 17, pp. 138-140, January, 1939.
107. Eine Quarzuhr für Zeit- und Frequenzmessung sehr hoher Genauigkeit. A. Scheibe und U. Adelsberger. *Physikalische Zeitschrift*, v. 33, No. 21, pp. 835-841, November, 1932.
108. Die Technischen Einrichtungen der Quarzuhr der Physikalisch-Technischen Reichsanstalt. A. Scheibe und U. Adelsberger. *Hochfrequenztechnik und Elektroakustik*, v. 43, pp. 37-47, February, 1934.
109. Über die ersten Erfahrungen mit den Quarzuhr des Preussischen Geodätischen Instituts, E. Kohlschütter. Verhandlungen der Siebenten Tagung der Baltischen Geodätischen Kommission. 1935.
110. German Quartz Clocks. *B.I.O.S. Report No.* 1316. H. M. Stationery Office, London. See also *Science Abstracts B*, page 242, September, 1947, and *Electrical Engineering Abstracts B*, September, 1947.
111. Quartzuhr und Normalfrequenz-Generator, L. Rhode und R. Leonhardt, E. N. T., pp. 117-122, June, 1940.
112. The national primary standard of radio frequency, E. L. Hall, V. E. Heaton and E. G. Clapham. *J. Research of National Bureau of Standards*, v. 14, pp. 85-98, 1935.
113. Crystal clock for accurate time standard, Vincent E. Heaton, *Instruments*, v. 20, No. 7, pp. 618-619 July, 1947.
114. W.W.V. standard frequency broadcasts, W. D. George, *F.M. and Television*, v. 7, pp. 25-27, June, 1947.
115. Time service of the U. S. Naval Observatory, J. Frederick Hellweg, *Trans. A. I. E. E.*, v. 51, No. 2, pp. 538-540, June, 1932.
116. Quartz crystal clocks—How Greenwich mean time is determined, Humphry M. Smith, *Electrical Times*, London, pp. 448-451, March 28, 1946.
117. The Measurement of Time, H. Spencer Jones. Reports on Progress in Physics, v. 4, pp. 1-26, 1937. The University Press, Cambridge, 1938.
118. Proceedings of Observatories—Royal Observatory, Greenwich. H. Spencer Jones, *Monthly Notices of R.A.S.*, v. 103, No. 2, pp. 77, 78, 1943.
119. The short-period erratics of free pendulum and quartz clocks, W. M. H. Greaves and L. S. T. Symms, *Monthly Notices of R. S. A.*, v. 103, No. 4, pp. 196-209, 1943.
120. (a) High precision timekeeping—Quartz clock now superseded free pendulum. F. Hope-Jones, *Horological Journal*, October, 1943.
120. (b) The precision timekeeping revolution—More about the Quartz Clock. F. Hope-Jones, *Horological Journal*, November, 1943.
121. A quartz clock, C. F. Booth, *P.O.E.E. Journal*, v. 39, Part II, pp. 33-37, July, 1946.
122. A standard of frequency and its applications, C. F. Booth and F. J. M. Laver, *The Journal of the I. E. E.*, v. 93, part III, pp. 223-236, July, 1946.
123. Theory and practice of pendulum observations at sea, F. A. Vening Meinesz. Netherlands Geodetic Commission, 1929.
124. Gravitational force measuring apparatus, W. A. Marrison. *U. S. Patent No.* 2,319,940, filed Sept., 1939, issued May, 1943.
125. The Dye quartz ring as a standard of frequency and time, L. Essen, *Proc. Royal Soc.*, v. 155A, pp. 498-519, July, 1936.
126. A new form of frequency and time standard, L. Essen, *Proc. Phys. Soc. of London*, v. 50, pp. 413-426, 1938.
127. Crystal oscillators for radio transmitters: An account of experimental work carried out by the Post Office. C. F. Booth and E. J. C. Dixon. *Jl. I. E. E.*, v. 77, pp. 197-236, 1935.
128. Thermostat design for frequency standards, W. A. Marrison, *Proc. I. R. E.*, v. 16, No. 7, pp. 976-980, July, 1928.
129. Superconductivity of lead at 3-cm. wavelength. F. Bitter, J. B. Garrison, J. Halpern, E. Maxwell, J. C. Slater and C. F. Squire, *Phys. Rev.*, v. 70, pp. 97, 98, July 1, 1946.
130. Enc. Brit. 14th Ed. See "Joseph von Fraunhofer".

131. Short Electric Waves, E. F. Nicols and J. D. Tear, *Phys. Rev.*, v. 21, pp. 587-610, June, 1923.
132. The shortest continuous waves, C. E. Cleeton and N. H. Williams, *Phys. Rev.*, v. 50, p. 1091, December, 1936.
133. A millimeter-wave reflex oscillator, J. M. Lafferty, *Jl. of Applied Physics*, v. 17, No. 12, pp. 1061-1066, December, 1946.
134. Electromagnetic waves of 1.1 cm. wavelength and the absorption spectrum of ammonia, C. E. Cleeton and N. H. Williams, *Phys. Rev.*, v. 45, pp. 234-237, Feb. 15, 1934.
135. The ammonia spectrum and line shapes near 1.25-cm. wavelength, Charles Hard Townes, *Phys. Rev.*, v. 70, pp. 665-671, November, 1946.
136. The inversion spectrum of ammonia, William E. Good, *Phys. Rev.*, v. 70, pp. 213-218, August, 1946.
137. A new method for measuring nuclear magnetic moments. (Letter to the editor), I. I. Rabi, J. R. Zacharias, S. Millman and P. Kusch, *Phys. Rev.*, v. 53, p. 318, 1938.
138. Molecular beam technique, I. Estermann. *Reviews of Modern Physics*, v. 18, No. 3, pp. 300-323, July, 1946.
139. The molecular beam magnetic resonance method. The radiofrequency spectra of atoms and molecules. J. B. M. Kellogg and S. Millman, *Reviews of Modern Physics*, v. 18, No. 3, pp. 323-352, July, 1946.
140. The rotation of the earth, and the secular accelerations of the sun, moon and planets. H. Spencer Jones, Astronomer Royal, *Monthly Notices of R. A. S.*, v. 99, pp. 541-558, May, 1939.

Abstracts of Technical Articles by Bell System Authors

*Experimental Determination of Helical-Wave Properties.*¹ C. C. CUTLER. The properties of the wave propagated along a helix used in the traveling-wave amplifier are discussed. A description is given of measurements of field strength on the axis, field distribution around the helix, and the velocity of propagation. It is concluded that the actual field in the helix described is slightly weaker than would be predicted from the relations presented by J. R. Pierce for a hypothetical helical surface.

*Results of Microwave Propagation Tests on a 40-Mile Overland Path.*² A. L. DURKEE. This paper gives the results of a series of microwave radio propagation tests over an unobstructed 40-mile overland path. The purpose of the tests was to investigate the transmission characteristics of such a path at centimeter wavelengths over a long period of time. Statistics on the transmission results at wavelengths ranging from 1.25 to 42 cm. are given. The tests extended over a period of about two years.

*A Tunable Vacuum-Contained Triode Oscillator for Pulse Service.*³ C. E. FAY* and J. E. WOLFE. A tunable push-pull triode oscillator is described in which the vacuum-tube components and the entire r.f. portion of the oscillator circuit are contained in an evacuated metallic envelope. A terminal is provided for coaxial output into a 50-ohm transmission line. The oscillator was developed for the frequency range of 390 to 435 Mc. and is tunable by mechanical means continuously through this range. Pulse power of above $\frac{1}{2}$ megawatt is obtained with pulse voltages of 15 to 17 kilovolts applied.

*A Proposed Loudness-Efficiency Rating for Loudspeakers and the Determination of System Power Requirements for Enclosures.*⁴ H. F. HOPKINS and N. R. STRYKER. Experimental and computed data relating to the loudness contribution of various ranges of the frequency spectra of speech and music are correlated with the corresponding energy distribution. A relatively simple measurement of sound pressure and a knowledge of certain acoustic radiation phenomena are applied to this correlation to form the basis of a

¹ *Proc. I. R. E.*, February 1948.

² *Proc. I. R. E.*, February 1948.

³ *Proc. I. R. E.*, February 1948.

* Of Bell Tel. Labs.

⁴ *Proc. I. R. E.*, March 1948.

method for predicting the loudness established by loudspeakers in enclosures. A loudness-efficiency rating for loudspeakers is suggested, and its application to sound-system engineering problems is described.

*A Sheet of Air Bubbles as an Acoustic Screen for Underwater Noise.*⁵ DONALD P. LOYE* and WM. FRED ARNDT. In Pearl Harbor, where there often were eight hundred ships of all kinds, the underwater noise level was high. No place was found where noise measurements could be made satisfactorily, and therefore it was decided that the best arrangement would be to insulate Auxiliary Repair Docks and measure the noise of submarines while they were in the docks. This was done by the development of a suitable air bubble screen across the open end of the dock. Such an acoustic barrier was comparatively easy to install, did not interfere with submarines entering and leaving, kept ocean surface oil out of the dock, insulated against low- as well as high-frequency noises as was required and, after extensive experimentation, the noise of the screen was reduced to a level that did not interfere with the noise measurements. The insulation of the screen upon the noise of a nearby submarine charging batteries is illustrated by a phonograph recording.

*A Method of Determining and Monitoring Power and Impedance at High Frequencies.*⁶ J. F. MORRISON and E. L. YOUNKER. A method and newly developed devices for determining and monitoring power and impedance levels in transmission lines at high frequencies are explained. Practical considerations influencing accurate determination of power and impedance levels are analyzed, and the previous and newly developed methods of monitoring these important quantities under changing conditions of load are compared.

*Automatic Volume Control as a Feedback Problem.*⁷ B. M. OLIVER. Feedback amplifier theory is shown to be applicable to the usual a.v.c. system. Expressions are derived for the loop gain in terms of the design requirements and the gain-control characteristic of the controlled amplifier. Using these expressions, the design of an a.v.c. system is quite straightforward and its characteristics, such as regulation and effect on desired modulation, are readily predictable.

⁵ *Jour. Acous. Soc. Amer.*, March 1948.

* Of Western Electric Co.

⁶ *Proc. I. R. E.*, February 1948.

⁷ *Proc. I. R. E.*, April 1948.

Contributors to this Issue

W. R. BENNETT, B.S., Oregon State College, 1925; A.M., Columbia University, 1928. Bell Telephone Laboratories, 1925-. Mr. Bennett has been active in the design and testing of multichannel communication systems, particularly with regard to modulation processes and the effects of nonlinear distortion. As a member of the Transmission Research Department, he is now engaged in the study of pulse modulation techniques for sending telephone channels by microwave radio relay.

H. T. BUDENBOM, B.S.E.E., 1922 and E.E., 1928, Purdue University; part time postgraduate work at Columbia and New York Universities. Western Electric Company, Engineering Department, and Bell Telephone Laboratories since 1922. Wire transmission problems, radio interference studies, point-to-point communication facilities and aviation radio development until 1933. Since 1933 Mr. Budenbom's development work has been primarily on classified projects, which have included radio direction finder and various types of radars. His present activities remain primarily in the field of classified military electronics.

CHARLES CLOS, C.E., New York University, 1927. New York Telephone Company, plant extension engineering, valuation and depreciation matters, intercompany settlements and tandem and tool fundamental plans, 1927-47. Pratt Institute, Evening School, Mathematics Instructor, 1946-48. Bell Telephone Laboratories, studies on development planning for local and toll switching systems, 1947-.

WARREN A. MARRISON, B.Sc. in Physics, Queen's University, Kingston, Canada, 1920; M.A. in Mathematics and Physics, Harvard University, 1921. Engineering Department, Western Electric Company, 1921-25; Bell Telephone Laboratories, 1925-. Mr. Marrison's work has consisted largely of the development and applications of precise standards of frequency.

S. A. SCHELKUNOFF, B.A., M.A. in Mathematics, The State College of Washington, 1923; Ph.D. in Mathematics, Columbia University, 1928. Engineering Department, Western Electric Company, 1923-25; Bell Telephone Laboratories, 1925-26. Department of Mathematics, State College of

Washington, 1926-29. Bell Telephone Laboratories, 1929-. Dr. Schelkunoff has been engaged in mathematical research, especially in the field of electromagnetic theory.

CLAUDE E. SHANNON, B.S. in Electrical Engineering, University of Michigan, 1936; S.M. in Electrical Engineering and Ph.D. in Mathematics, M.I.T., 1940. National Research Fellow, 1940. Bell Telephone Laboratories, 1941-. Dr. Shannon has been engaged in mathematical research principally in the use of Boolean Algebra in switching, the theory of communication, and cryptography.

Public Library
Kansas City, Mo.

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS
OF ELECTRICAL COMMUNICATION

Equivalent Circuits of Linear Active Four-Terminal Net-
works *L. C. Peterson* 593

A Mathematical Theory of Communication (Concluded)
C. E. Shannon 623

Transients in Mechanical Systems *J. T. Muller* 657

Maximally-Flat Filters in Wave Guide .. *W. W. Mumford* 684

Transient Response of an FM Receiver *M. K. Zinn* 714

Transverse Fields in Traveling-Wave Tubes .. *J. R. Pierce* 732

Abstracts of Technical Articles by Bell System Authors... 747

Contributors to this Issue 752

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

THE BELL SYSTEM TECHNICAL JOURNAL

*Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York, N. Y.*



EDITORS

R. W. King

J. O. Perrine

EDITORIAL BOARD

C. F. Craig

O. E. Buckley

O. B. Blackwell

M. J. Kelly

H. S. Osborne

A. B. Clark

J. J. Pilliod

F. J. Feely



SUBSCRIPTIONS

Subscriptions are accepted at \$1.50 per year. Single copies are 50 cents each.
The foreign postage is 35 cents per year or 9 cents per copy.



Copyright, 1948
American Telephone and Telegraph Company

The Bell System Technical Journal

Vol. XXVII

October, 1948

No. 4

Equivalent Circuits of Linear Active Four-Terminal Networks*

By

LISS C. PETERSON

INTRODUCTION

THE art of equivalent network representation has grown very considerably since its inception by Dr. G. A. Campbell. In his paper "Cissoidal Oscillations" which was published in 1911 he proved that any passive network made up of a finite number of invariable elements and having one pair of input terminals and one pair of output terminals is externally equivalent to an unsymmetrical T or Π network. From this modest beginning the field of applications of the equivalent circuit concept has steadily expanded so that by now the whole field of linear passive circuit theory has been subjected to equivalent circuit interpretation.

With the advent of the thermionic vacuum tube amplifier, linear active network theory had to be considered and almost immediately the attempt was made to obtain an equivalent circuit whose performance would depict the linear characteristic of the tube. The equivalent circuit art has also, in recent years, been used to describe the performance of certain classes of non-linear devices, such as mixers, and further applications in this field will no doubt be made.

Equivalent circuit concepts have played an important part in electrical

* This paper appears substantially as it was originally prepared some years ago as a technical memorandum for internal distribution within Bell Telephone Laboratories. Its publication is rendered timely by the applicability to the recently announced transistor devices. Present experience indicates, in fact, that the configuration of Fig. 13 furnishes the most useful equivalent four-pole network for the transistor.

Mr. J. A. Morton has called my attention to an early paper in this field by Strecker and Feldtkeller (E.N.T. Vol. 6, page 93, 1929) in which the general theory of active networks has been well developed. However, the early state of the then prevailing art prevented the full demonstration of the power of the method and of course precluded the possibility of application to modern devices. This paper enlarges the general theory and makes logical applications of the method to modern devices. Since the appearance of the Strecker-Feldtkeller paper several other papers touching upon this subject have appeared. However, no attempt at giving a complete bibliography will be made, except to call attention to the contributions of Prof. M. J. O. Strutt, who also has adopted the four-pole point of view.

The I.R.E. Electron Tube Committee has adopted the four-pole viewpoint and has proposed methods of tests for experimentally determining the four-pole parameters of electron tubes. This material will be published in the new I.R.E. standards on electron tubes.

engineering, particularly in communication engineering. One might almost say that a problem is not solved unless an equivalent circuit has been found whose performance will exhibit some of the characteristics of the actual problem. This need for circuit concepts reflects a desire to make the phenomena more alive and subject to physical interpretation, for it is true that equivalent circuit concepts have greatly contributed towards physical interpretation of analytical expressions. A problem may very well be studied without recourse to the equivalent circuit concept and a correct answer obtained, yet the development of an equivalent circuit adds greatly to the complete interpretation of the physical phenomena.

In this paper we shall be concerned only with linear a-c amplifier operation where the term *linear* indicates that the analytical expressions connecting currents and voltages are linear, that is, involve only the first power of any instantaneous current or its derivative. We shall further restrict our attention to the usual mode of four-terminal amplifier operation in which one pair of terminals is associated with the signal to be amplified and the other pair with the amplified signal. The equivalent a-c circuit of such a transducer will require the development and interpretation of the linear relations connecting the a-c currents and voltages at the input terminals with corresponding quantities at the output terminals. At this point we can logically postulate that the important formal difference between an active and a passive four-pole lies in that the law of reciprocity no longer can be assumed to hold for the active network. Since passive four-poles require three independent parameters for their complete specification the active four-pole will require at least one additional parameter.

In the practical applications we shall be principally concerned with the various triode four-pole connections. A short review of the various stages involved in deriving the newer forms of equivalent triode circuits will therefore be considered prior to the introduction of generalized concepts. Such a review is in the main concerned with the effect of frequency upon the early forms of the equivalent triode circuit.

In the review we shall confine ourselves to the usual grounded cathode mode of operation since it is only in recent times that grounded grid and grounded plate operation have come into use. This distinction is, however, not necessary and is introduced solely for simplicity.

The conventional notation as well as the positive current directions are indicated on Fig. 1 for a general four-pole N . It is assumed here that terminals 1 are the available input and terminals 2 the available output terminals. This choice of current direction appears to the writer to be the most convenient to use when the four-pole is energized at the input terminals 1 only.

Consider, now, a grounded triode operated at such a low frequency that

all displacement or capacity currents can be disregarded, and let it first be supposed that the grid is negatively polarized with respect to the cathode so that grid current is absent. This represents the most primitive form of operation, governed by a law which has been termed "The equivalent-plate-circuit Theorem."¹ With reference to the assumed current direction this theorem may be expressed by saying that the application of the voltage V_1 to the grid is equivalent, so far as phenomena in the plate circuit are concerned, to the application of the voltage $-\mu V_1$ in series with a resistance r_p , where μ is the amplification factor and r_p the internal plate resistance. The

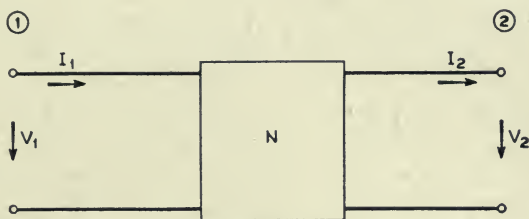


Fig. 1—Current-voltage relations for a general four-pole.

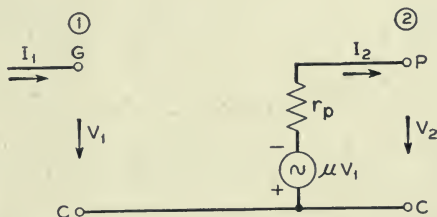


Fig. 2—Equivalent circuit of a negative grid triode at low frequencies.

equivalent circuit for this mode of operation is thus as shown in Fig. 2, where the input terminals 1 are represented by the grid G and the cathode C and the output terminals by the anode P and the cathode C. In terms of analysis the circuit is described by the two equations

$$\left. \begin{aligned} I_1 &= 0 \\ -\mu V_1 &= I_2 r_p + V_2 \end{aligned} \right\} \quad (1)$$

Equations (1) and their associated circuit Fig. 2 are expressions of the fact that in any closed loop the voltages must be in equilibrium.

By a slight rearrangement of (1) a network representation based on current

¹ Chaffee, "Theory of Thermionic Vacuum Tubes," page 192.

equilibrium may be obtained. For this purpose (1) is written as

$$\left. \begin{aligned} I_1 &= 0 \\ I_2 &= -\frac{\mu}{r_p} V_1 - \frac{1}{r_p} V_2 \end{aligned} \right\} \quad (2)$$

The corresponding network representation is as shown in Fig. 3, where the energizing source in the plate circuit consists of a constant current generator of strength $-\frac{\mu}{r_p} V_1$ impressed across the output terminals 2.

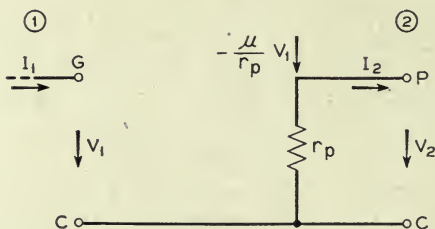


Fig. 3—Equivalent circuit of a negative grid triode at low frequencies.

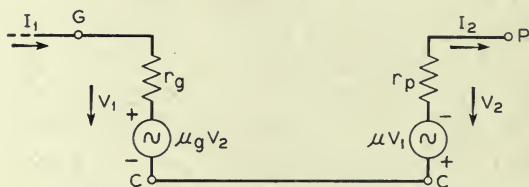


Fig. 4—Equivalent circuit of a positive grid triode at low frequencies.

Let us now consider a further step and assume that the grid is positive so that grid current is also flowing. In regard to the grid circuit there is a theorem called the "Equivalent-grid-circuit Theorem" which is exactly similar to the corresponding plate circuit theorem. The theorem says that the a-c grid current can be calculated by assuming that an e.m.f. $\mu_g V_2$ acts in series with a resistance r_g where μ_g is the reflex factor and r_g the internal grid resistance. In symbols and by using the notation of Fig. 1 this is expressed by:

$$V_1 = \mu_g V_2 + I_1 r_g \quad (3)$$

By combining the two theorems the equivalent circuit of Fig. 4 results. Again by writing (3) as

$$I_1 = \frac{1}{r_g} V_1 - \frac{\mu_g}{r_g} V_2 \quad (4)$$

a corresponding network based upon current equilibrium may be found.

It is well to note that this general low-frequency case is governed by the two simultaneous equations

$$\left. \begin{aligned} I_1 &= \frac{1}{r_g} V_1 - \frac{\mu_g}{r_g} V_2 \\ I_2 &= -\frac{\mu}{r_p} V_1 - \frac{1}{r_p} V_2 \end{aligned} \right\} \quad (5)$$

which are of the general form

$$\left. \begin{aligned} I_1 &= \beta_{11} V_1 + \beta_{12} V_2 \\ I_2 &= \beta_{21} V_1 + \beta_{22} V_2 \end{aligned} \right\} \quad (6)$$

where

$$\left. \begin{aligned} \beta_{11} &= \frac{1}{r_g} & \beta_{12} &= -\frac{\mu_g}{r_g} \\ \beta_{21} &= -\frac{\mu}{r_p} & \beta_{22} &= -\frac{1}{r_p} \end{aligned} \right\} \quad (7)$$

The network of Fig. 4 is thus a possible form of circuit interpretation of (5) or (6). It may also be observed that (6) represents at least formally the most general formulation of the linear active four-pole so that even at very low frequencies the general four-pole point of view might be useful.

Several observations may now be made. In the first place it should be noted that these networks are not based on any study of the internal action of the tube, but rather on the purely formal mathematical process of differentiating the two functional relations which express the broad fact that plate and grid currents are some unspecified linear continuous functions of the grid and plate potentials in the neighborhood of the operating point.

In the second place it may be observed that the network of Fig. 4 represents in a sense two separate networks interacting with each other by means of voltage or current generators. This method of equivalent circuit representation is the result of separate interpretation of the equivalent plate and grid circuit theorems. As a corollary it follows that such a four-pole equivalence involves at least two generators within the network in order to take the effect of interaction into account.

We may say that the networks discussed were satisfactory so long as the frequency was low enough to allow displacement currents to be disregarded. With the operation of circuits at higher frequencies (up to the order of 10^6 cps, say) it became necessary to take the internal tube capacitances into account. This was done by the superposition of a capacity network as shown in Fig. 5. It is interesting as well as instructive to formulate this network transi-

tion in analytical terms. The transition rests upon the physical fact that the total current entering or leaving an electrode is the sum of conduction and

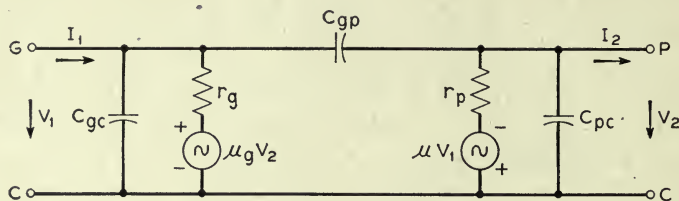


Fig. 5—Equivalent circuit of a positive grid triode at moderately low frequencies.

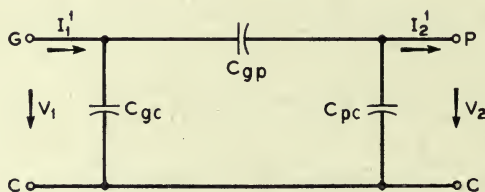


Fig. 6—Parasitic triode network.

displacement current. The network for the displacement currents is passive and is shown on Fig. 6. For this network:

$$\left. \begin{aligned} I_1' &= \beta_{11}' V_1 + \beta_{12}' V_2 \\ I_2' &= -\beta_{12}' V_1 + \beta_{22}' V_2 \end{aligned} \right\} \quad (8)$$

where I_1' is input and I_2' the output displacement current. The coefficients appearing have the values

$$\left. \begin{aligned} \beta_{11}' &= i\omega(C_{gc} + C_{gp}) \\ \beta_{12}' &= -i\omega C_{gp} \\ \beta_{22}' &= -i\omega(C_{pc} + C_{gp}) \end{aligned} \right\} \quad (9)$$

The potentials appearing in (8) are the same as those which occur in the conduction current equations (5), this being the physical condition which also must be satisfied. By adding (5) and (8) and by letting I_1 and I_2 mean total currents we get:

$$\left. \begin{aligned} I_1 &= \left[\frac{1}{r_g} + i\omega(C_{gc} + C_{gp}) \right] V_1 - \left[\frac{\mu_g}{r_g} + i\omega C_{gp} \right] V_2 \\ I_2 &= \left[-\frac{\mu}{r_p} + i\omega C_{gp} \right] V_1 + \left[\frac{1}{r_p} + i\omega(C_{pc} + C_{gp}) \right] V_2 \end{aligned} \right\} \quad (10)$$

These equations are again of the form

$$\left. \begin{aligned} I_1 &= \beta_{11} V_1 + \beta_{12} V_2 \\ I_2 &= \beta_{21} V_1 + \beta_{22} V_2 \end{aligned} \right\} \quad (11)$$

where now

$$\left. \begin{aligned} \beta_{11} &= \frac{1}{r_g} + i\omega(C_{gc} + C_{gp}) \\ \beta_{12} &= -\left[\frac{\mu_g}{r_g} + i\omega C_{gp}\right] \\ \beta_{21} &= -\frac{\mu}{r_p} + i\omega C_{gp} \\ \beta_{22} &= -\left[\frac{i}{r_p} + i\omega(C_{pc} + C_{gp})\right] \end{aligned} \right\} \quad (12)$$

The network interpretation of Fig. 5 is consistent with (10) and so does, in fact, constitute a possible network representation.

From (12) the relation

$$\beta_{21} + \beta_{12} = -\left[\frac{\mu}{r_g} + \frac{\mu_g}{r_g}\right] \quad (13)$$

is obtained. The thing to emphasize is that this sum is independent of the passive feedback admittance and that it is a constant independent of frequency within the range considered. It is, moreover, a quantity which only is correlated with the conduction currents in the tube. For lack of a better name the sum (13) will be referred to as "the effective transconductance." It will play an important role in the later discussion.

As still higher frequencies (above 10^7 cps) were employed it became necessary to take into account lead effects, usually in the form of self and mutual inductances, and by incorporating them in the network of Fig. 5 a slightly more involved circuit was obtained. This more general network is still determined by equations of the form (11), for adding the new network elements merely means that linear transformations are applied to the potentials V_1 and V_2 with the result that a new set of β -coefficients is obtained, the new set being merely of somewhat more complicated form than the first.

With the utilization of frequencies so high that the electron transit times became comparable with the period of the applied signal, further complications arose and the equivalent network idea was put to a severe test. In

this development we may distinguish between two methods of approach. In one the attempt was made to modify the low-frequency network of Fig. 5 to include transit time effects to a first order of approximation.² The second approach differed in that attention was directed only toward the electron stream itself, while the circuit elements connecting the stream with the physically available terminals were grouped together with the external circuit elements.³ The latter approach represented a particularly useful one from a physical point of view and it also extended the use of basic circuit elements to include the general diode impedance as a new circuit element complete in itself. However, even in this latest approach the four-pole point of view was not adopted, with consequent loss of generality and unity in viewpoint. Moreover, the fact that only the electron stream itself was considered caused some confusion.

With this brief review of the development of equivalent circuit representation of vacuum tube amplifiers in mind we turn now to the main body of the paper in which a more general treatment of the problem is considered. It will be shown, in the coming sections, how it is possible to lump all the factors involved in vacuum tube amplifiers, i.e., physical circuit parameter and internal electronic effects involving the electron transit time, into a single coordinated picture with an equivalent circuit representation of the overall effect.

EQUIVALENT CIRCUIT REPRESENTATION OF ACTIVE LINEAR FOUR-POLE POLE EQUATIONS

Whenever the response of a general transducer is related in a linear manner to the stimulus, the transducer behavior is described by two linear relations. Although we are primarily concerned with electromagnetic transducers the concepts to be used are of broader utility and may, for example, also be applied to mechanical and electromechanical transducers.

There are various ways in which the behavior of the four-pole may be expressed analytically. The form expressing current equilibrium has already been given and it may well serve as a starting point for the following discussion:

Thus we have:

$$\left. \begin{aligned} I_1 &= \beta_{11}V_1 + \beta_{12}V_2 \\ I_2 &= \beta_{21}V_1 + \beta_{22}V_2 \end{aligned} \right\} \quad (14)$$

² F. B. Llewellyn, "Electron-Inertia Effects," Cambridge University Press, 1941.

³ F. B. Llewellyn and L. C. Peterson, Interpretation of Ultra-High Frequency Tube Performance in Terms of Equivalent Networks, Proceedings of the National Electronics Conference, 1944.

as expressions for current equilibrium at any frequency. The four parameters β which appear have simple physical meanings, and it is seen that

β_{11} is the input admittance with output shorted

$-\beta_{22}$ is the output admittance with input shorted

$-\beta_{12}$ is the feedback admittance with input shorted

β_{21} is the transfer admittance with output shorted.

Before proceeding to the network representations of (14) it seems well to state briefly some of the reasons which almost force one to adopt the four-pole point of view when dealing with vacuum tubes in the higher frequency range.

There are the pedagogical reasons that classical methods long employed for passive networks are merely extended into the realm of active networks, thus providing unity in viewpoints.

The basic analysis involving the four-pole parameters for a particular transducer needs to be performed only once and, once obtained, all problems involving terminal impedances may, in any particular case, be solved in a routine manner.

There are also further practical reasons. We saw above that with increase in frequency the classical equivalent network had to be modified to a considerable extent in order to include the parasitic elements. This poses a serious problem for the tube designer, whose task it is to estimate the tube performance between known terminations. Such a task based upon the modified classical circuit becomes very difficult and cumbersome. Moreover, it is also difficult to segregate and measure the parasitic elements. Hence it appears that one could gain much if design parameters could be developed capable of reflecting parasitic and transit time effects. Finally, it is desirable to develop equivalent circuits with a minimum number of parameters bearing simple relationships to quantities which can be measured directly.

These general desires arising from the practical needs of the tube designer can be satisfied if tube behavior is specified by means of four-pole parameters.

All in all the four-pole point of view can be made to satisfy the logical needs of integrated concepts as well as the practical needs of simplicity in the specification of tube performance.

After this brief presentation of the argument for the four-pole point of view, the network representation of (14) will now be considered.

Stated in broadest terms: we are seeking a network representation by considering the two equations as a single unit and not by the trivial consideration of each equation by itself.

We need, to begin with, the well-known network representation of a passive four-pole. Equation (14) has, then, the form

$$\left. \begin{aligned} I_1 &= \beta_{11}V_1 + \beta_{12}V_2 \\ I_2 &= -\beta_{12}V_1 + \beta_{22}V_2 \end{aligned} \right\} \quad (15)$$

and one equivalent circuit representation is that given by the Π network having the element values shown in Fig. 7. If so desired the Π network can of course also be transferred into an equivalent T network.

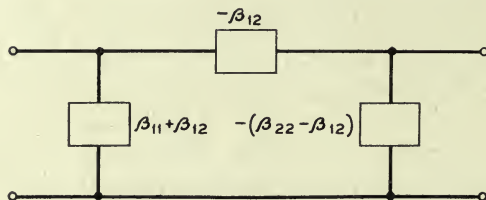


Fig. 7—Equivalent circuit of a passive four-pole.

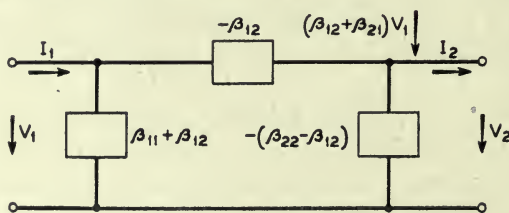


Fig. 8—Equivalent circuit of an active four-pole; current impressed at the output.

Now write (14) as

$$\left. \begin{aligned} I_1 &= \beta_{11}V_1 + \beta_{12}V_2 \\ I_2 &= -\beta_{12}V_1 + \beta_{22}V_2 + (\beta_{12} + \beta_{21})V_1 \end{aligned} \right\} \quad (16)$$

Whence it is seen by a comparison with (15) that the network representation of the active four-pole differs from the passive one merely by the presence of the impressed current $(\beta_{12} + \beta_{21})V_1$. A possible network representation of the general active four-pole is thus as shown on Fig. 8.

An immediate application may be illuminating. Consider, for example, the triode operated with positive grid, with interelectrode capacitances taken into account. The four-pole equations are given by (11) and (12) and the classical network is that of Fig. 5. From (12) and Fig. 8 we get the network of Fig. 9. It may be observed that, while in Fig. 5 the source and source-free constituents are intermingled, this is not the case in Fig. 9 where, on the contrary, a clear demarcation is present between such constituents.

In regard to the general network of Fig. 8 it may be noted that the network is composed of two parts. One part obeys the reciprocal law and is represented by a Π (or T) network and is consequently specified by three parameters. The other part is merely an impressed current controlled by the input potential V_1 . From the fact that the Π (or T) network obeys the

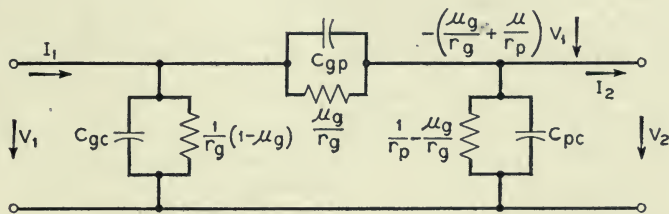


Fig. 9—Equivalent circuit of a positive grid triode at moderately low frequencies; current impressed at the output.

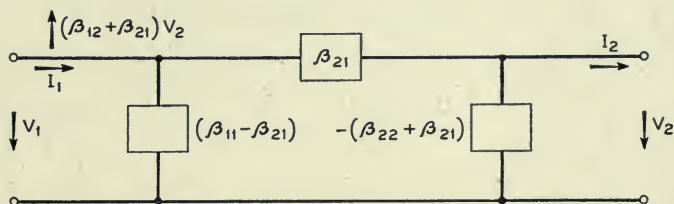


Fig. 10—Equivalent circuit of an active four-pole; current impressed at the input.

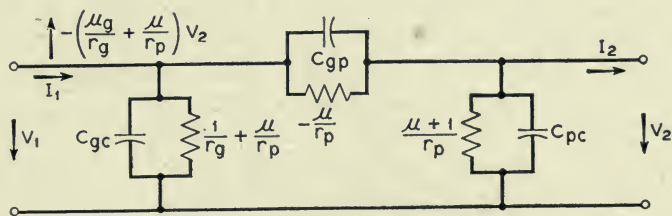


Fig. 11—Equivalent circuit of a positive grid triode at moderately low frequencies; current impressed at the input.

reciprocal law, the conclusion does not follow that it in general behaves as a passive network. The element values may, for example, be negative, and Bode's integral relations may not necessarily be true for this network.

It should further be noted that the network representation holds for all frequencies. The effect of frequency will be that the admittances in the network change and in changing they will, for example, reflect effects due to parasitics and electron transit time.

We next observe that the network of Fig. 8 is not unique; in other words it is not the only possible network. To see this it need merely be observed that from the group of four independent β parameters in (15) there are several ways in which a subgroup may be selected containing only three of them. For example the "passive part" of the network in Fig. 8 was constructed from the three parameters β_{11} , β_{12} and β_{22} . But this is evidently not the only choice. Moreover, the impressed force was taken to be voltage-controlled but this again does not represent the only possibility.

In general it follows that the "passive part" of the network will reflect at least three properties of the complete network. In Fig. 8, for example, the "passive part" reproduces faithfully the two short-circuit driving-point admittances and the feedback admittance of the complete network.

Finally it is well to observe that only one driving force is needed in the general network formulation.

With this background of the general ideas involved let us now further explore some of the possibilities suggested as to other "passive network" constituents. Let us, for example, use β_{11} , β_{22} and β_{21} for the "passive part". We then write (14) as

$$\left. \begin{aligned} I_1 &= \beta_{11}V_1 - \beta_{21}V_2 + (\beta_{12} + \beta_{21})V_2 \\ I_2 &= \beta_{21}V_1 + \beta_{22}V_2 \end{aligned} \right\} \quad (17)$$

and from (17) the network representation of Fig. 10 follows. In addition this network differs from that of Fig. 8, in that the impressed current appears on the input side and that it is controlled by the output rather than by the input voltage. As an illustration consider again the triode operated with positive grid. The equivalent network is now as shown in Fig. 11.

Now let it be supposed that the impressed force be current rather than voltage-controlled. We then first transform (14) into

$$\left. \begin{aligned} V_1 &= \frac{I_1}{\beta_{11}} - \frac{\beta_{12}}{\beta_{11}} V_2 \\ I_2 &= \frac{\beta_{21}}{\beta_{11}} I_1 + V_2 \left(\beta_{22} - \frac{\beta_{21}\beta_{12}}{\beta_{11}} \right) \end{aligned} \right\} \quad (18)$$

and then rewrite (18) as

$$\left. \begin{aligned} V_1 &= \frac{I_1}{\beta_{11}} - \frac{\beta_{12}}{\beta_{11}} V_2 \\ I_2 &= -\frac{\beta_{12}}{\beta_{11}} I_1 + V_2 \left(\beta_{22} - \frac{\beta_{21}\beta_{12}}{\beta_{11}} \right) + \frac{\beta_{12} + \beta_{21}}{\beta_{11}} I_1 \end{aligned} \right\} \quad (19)$$

It can be shown that the network representation of Fig. 12 is consistent with (19). This network representation suffers from two obvious disadvantages. One is that the "passive network" is determined by only one short-circuit driving-point admittance of the complete network. The other short-circuit driving-point admittance of the "passive part" appears to be unrelated to any simple admittance which may be found from measurements at the output terminals of the complete network. The second disadvantage is that the coefficient of the current in the impressed force is of a complicated nature, since a driving-point admittance enters. Moreover, a closer investigation shows that in this network representation the "passive part" besides preserving the driving-point admittance β_{11} and the feedback admittance β_{12} , also preserves the quantity $\Delta_\beta = \beta_{11}\beta_{22} - \beta_{12}\beta_{21}$ of the complete network.

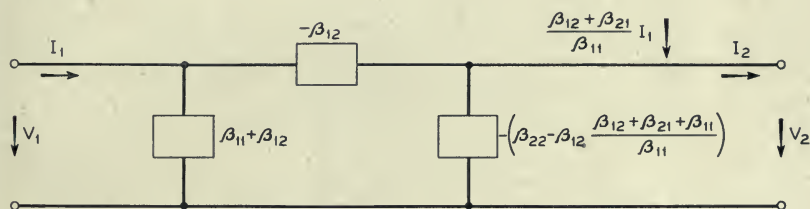


Fig. 12—Equivalent circuit of an active four-pole; current impressed at the output.

By proceeding from (19) in a way similar to that used in (17) the network in Fig. 12 transforms into another in which the impressed force appears on the input side. Many other networks can also be found.

The networks discussed were based on (14), which expressed the fact of current equilibrium and leads rather naturally to Π networks together with an impressed current source. On the other hand, starting with the four-pole equations which express voltage equilibrium, one encounters T networks together with impressed electromotive forces. These networks must now also be considered. In regard to the details involved in their derivation we may be very brief since the methods are similar to those already employed.

The four-pole equations expressing voltage equilibrium may be written as

$$\left. \begin{aligned} V_1 &= Z_{11}I_1 + Z_{12}I_2 \\ V_2 &= Z_{21}I_1 + Z_{22}I_2 \end{aligned} \right\} \quad (20)$$

where current and voltage directions are assumed to be taken in accordance with the conventions of Fig. 1.

The four parameters Z are of simple physical significance and it follows that:

- Z_{11} is the input impedance with output open
- $-Z_{22}$ is the output impedance with input open
- $-Z_{12}$ is the feedback impedance with input open
- Z_{21} is the transfer impedance with output open.

The relations between the β 's and the Z 's are given by the expressions:

$$\left. \begin{aligned} \beta_{11} &= \frac{Z_{22}}{\Delta_Z} \\ \beta_{12} &= -\frac{Z_{12}}{\Delta_Z} \\ \beta_{21} &= -\frac{Z_{21}}{\Delta_Z} \\ \beta_{22} &= \frac{Z_{11}}{\Delta_Z} \end{aligned} \right\} \quad (21)$$

where

$$\Delta_Z = \begin{vmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{vmatrix} \quad (22)$$

We have also

$$\Delta_Z = \frac{1}{\Delta_\beta} \quad (23)$$

where

$$\Delta_\beta = \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{vmatrix} \quad (24)$$

Applying now to (20) the transformations which led to the networks of Figs. 9, 10 and 12, we get the networks of Figs. 13, 14 and 15.

The "passive part" of the network in Fig. 13 reproduces the two open circuit impedances Z_{11} and Z_{22} as well as the feedback impedance Z_{12} of the complete network.

The network in Fig. 14 differs from that of Fig. 13 because of the use of the transfer impedance Z_{21} in the "passive part" with the result that the impressed current appears on the input side.

The "passive part" in Fig. 15, finally, preserves the open-circuit impedance Z_{11} , the feedback impedance Z_{12} and the determinant Δ_z of the complete network. This network shows incidentally a close resemblance to one already published.³

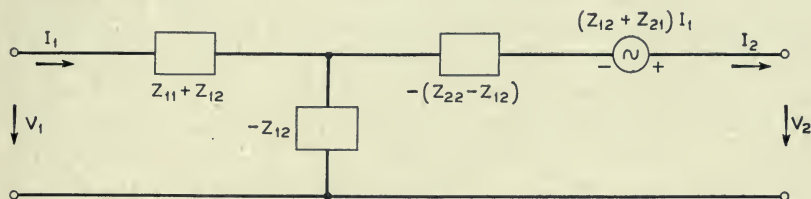


Fig. 13—Equivalent circuit of an active four-pole; voltage impressed in series with the output.

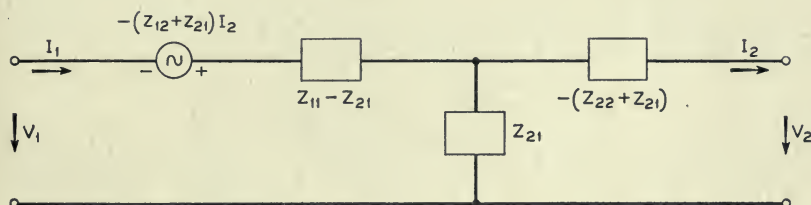


Fig. 14—Equivalent circuit of an active four-pole; voltage impressed in series with the input.

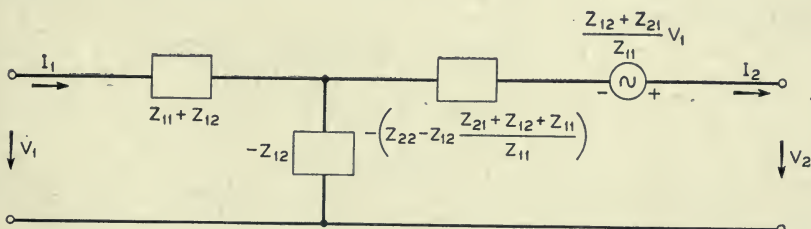


Fig. 15—Equivalent circuit of an active four-pole; voltage impressed in series with the output.

It is well to emphasize that, while all the complete networks are equivalent, this is not true for the "passive parts." In fact from their derivation it follows that none are equivalent. Specific circumstances may make it desirable to perform transformations on the passive parts. For example, it might be more convenient to work with a Π than with a T network. Such transformations are of course perfectly legitimate and raise the question of choice of equivalent networks. A few remarks on this subject may be

³ Loc. cit.

appropriate, since the choice is not quite a matter of indifference. Broadly speaking, the choice will depend upon the relative advantages of nodal and mesh analysis and in most practical situations the former has proved to be the more convenient of the two. One cannot, however, be too dogmatic in this regard. Consider the networks of Figs. 8 and 13. Suppose, for example, that parasitic elements appearing as a passive Π network had to be superimposed. It is then more convenient to use Fig. 8, not only on account of the ease with which this may be done, but also on account of the fact that the effective transadmittance $\beta_{12} + \beta_{21}$ is invariant with respect to such a superposition. If, on the other hand, parasitic elements appear as series elements (lead inductances for example) the network of Fig. 13 might be more convenient since the effective transimpedance $Z_{12} + Z_{21}$ now remains invariant.

It is also desirable to choose an equivalent network whose elements are capable of being determined by simple measurements; from this consideration the network on Fig. 8 is of distinct advantage.

APPLICATION TO TRIODES

The preceding section was primarily directed towards the development of possible forms of network representations of the general four-pole equations. In this section one of these forms, namely that given in Fig. 8, will be used to represent the three modes of triode operation. Depending upon which electrode is at a-c ground potential, we may distinguish between the following methods of operation:

1. Grounded cathode operation.
2. Grounded grid operation.
3. Grounded plate operation.

The schematic diagrams, together with assumed voltage and current directions for these modes, are shown on Figs. 16, 17 and 18 respectively.

With a given set of available terminals the first step in obtaining the networks consists in calculating the four-pole parameters with respect to these terminals. It will be assumed that the coupling circuits have been designed with such efficiency that lead effects can be disregarded, so that the available terminals actually coincide with anode, grid and cathode. This set of available terminals brings us as close to the electron stream as it is physically possible to attain and it represents the ideal towards which design tends.

It is beyond the scope of this paper to consider the details involved in the calculations of the four-pole parameters. The basic tools needed are the result of a study of the dynamics of the electron stream, which started from

fundamentals,⁴ and some familiarity with this work is assumed on the part of the reader. Concerning these tools two reservations need be made. In the first place the tools apply to planar rather than to cylindrical structures. Since, however, there is a decided tendency toward planar structures, especially in the high-frequency field, because of a desire for uniform electron streams, this limitation does not seem serious. In the second place the tools are also subject to the limitation of a single-valued velocity electron stream.

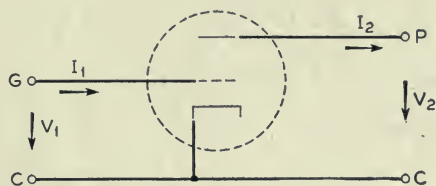


Fig. 16—Current-voltage relations for the grounded cathode triode.

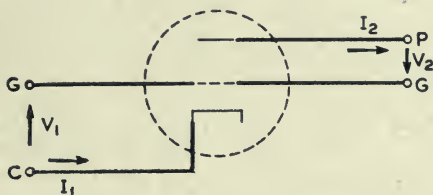


Fig. 17—Current-voltage relations for the grounded grid triode.

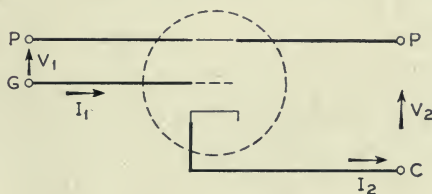


Fig. 18—Current-voltage relations for the grounded plate triode.

This, again, is not too serious, since one of the aims in present high-frequency tube design is to produce as uniform a stream as possible. Nevertheless, the effects produced by multiple velocities are important to know. Studies along such lines have been made by Mr. Frank Gray of these Laboratories.

The operating conditions of the triode are assumed to be quite general. There are, for example, no restrictions placed upon frequency and space charge and the grid may, moreover, have either positive or negative d-c potential with respect to the cathode.

⁴ F. B. Llewellyn and L. C. Peterson, "Vacuum Tube Networks," Proceedings of the I.R.E., March, 1944.

With current and voltage directions as in Figs. 16, 17 and 18, the following Tables I, II and III list the four-pole parameters for the three modes of triode operation in both β and Z forms.

TABLE I
FOUR-POLE PARAMETERS FOR GROUNDED CATHODE TRIODE

$$\left. \begin{aligned}
 \beta_{11} &= \frac{y_{11} + y_{21} + y_{22}}{D} & \beta_{12} &= -\frac{y_{22}}{D} \\
 \beta_{21} &= \frac{y_{21} + y_{22}}{D} & \beta_{22} &= -\frac{y_{22} + \frac{y_{11}}{\mu}}{D} \\
 \Delta_{\beta} &= \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{vmatrix} = -\frac{y_{11} y_{22}}{D} \\
 D &= 1 + \frac{y_{11} + y_{21} + y_{22}}{\mu y_{22}}
 \end{aligned} \right\}$$

$$\left. \begin{aligned}
 Z_{11} &= \frac{1}{y_{11}} + \frac{1}{\mu y_{22}} & Z_{12} &= -\frac{1}{y_{11}} \\
 Z_{21} &= \frac{1}{y_{11}} + \frac{y_{21}}{y_{11} y_{22}} & Z_{22} &= -\left[\frac{1}{y_{11}} + \frac{1}{y_{22}} + \frac{y_{21}}{y_{11} y_{22}} \right] \\
 \Delta_Z &= \begin{vmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{vmatrix} = \frac{1}{\Delta_{\beta}} = -\frac{D}{y_{11} y_{22}}
 \end{aligned} \right\}$$

μ = amplification factor.

The y admittance coefficients appearing in the above tables were fully explained and discussed in the paper on Vacuum Tube Networks, to which reference has already been made. Suffice it here to say that y_{11} is the admittance of the diode coinciding with cathode and equivalent grid plane and y_{22} the admittance of the diode coinciding with the equivalent grid plane and the anode and finally y_{21} the transadmittance between these fictitious diodes. The admittance y_{11} depends upon the d-c conditions between cathode and grid and upon the transit angle for this region alone. The diode admittance y_{22} depends in a similar manner upon the d-c space charge conditions in the grid-anode region as well as upon the transit angle for this region alone. For the small degree of space charge which usually exists between grid and plate of most triodes, y_{22} can be represented by a simple capacitance. The transadmittance y_{21} can be resolved into two factors, the first of which depends only upon the transit angle between cathode and grid, and the second only upon the transit angle between grid and anode. In the paper on vacuum tube networks all these admittances were plotted

TABLE II
FOUR-POLE PARAMETERS FOR GROUNDED GRID TRIODE

$$\left. \begin{aligned}
 \beta_{11} &= \frac{y_{11} \left(1 + \frac{1}{\mu} \right)}{D} & \beta_{12} &= -\frac{y_{11}}{\mu D} \\
 \beta_{21} &= -\frac{y_{21} - \frac{y_{11}}{\mu}}{D} & \beta_{22} &= -\frac{y_{22} + \frac{y_{11}}{\mu}}{D} \\
 \Delta_{\beta} &= \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{vmatrix} = -\frac{y_{11} y_{22}}{D} \\
 D &= 1 + \frac{y_{11} + y_{21} + y_{22}}{\mu y_{22}}
 \end{aligned} \right\}$$

$$\left. \begin{aligned}
 Z_{11} &= \frac{1}{y_{11}} + \frac{1}{\mu y_{22}} & Z_{12} &= -\frac{1}{\mu y_{22}} \\
 Z_{21} &= \frac{1}{\mu y_{22}} - \frac{y_{22}}{y_{11} y_{22}} & Z_{22} &= -\left[\frac{1}{y_{22}} + \frac{1}{\mu y_{22}} \right] \\
 \Delta_Z &= \begin{vmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{vmatrix} = \frac{1}{\Delta_{\beta}} = -\frac{D}{y_{11} y_{22}}
 \end{aligned} \right\}$$

μ = amplification factor.

TABLE III
FOUR-POLE PARAMETERS FOR GROUNDED PLATE TRIODE

$$\left. \begin{aligned}
 \beta_{11} &= \frac{y_{11} + y_{21} + y_{22}}{D} & \beta_{12} &= -\frac{y_{11} + y_{21}}{D} \\
 \beta_{21} &= \frac{y_{11}}{D} & \beta_{22} &= -\frac{y_{11} \left(1 + \frac{1}{\mu} \right)}{D} \\
 \Delta_{\beta} &= \begin{vmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{vmatrix} = -\frac{y_{11} y_{22}}{D} \\
 D &= 1 + \frac{y_{11} + y_{21} + y_{22}}{\mu y_{22}}
 \end{aligned} \right\}$$

$$\left. \begin{aligned}
 Z_{11} &= \frac{1}{y_{22}} + \frac{1}{\mu y_{22}} & Z_{12} &= -\left[\frac{1}{y_{21}} + \frac{y_{21}}{y_{11} y_{22}} \right] \\
 Z_{21} &= \frac{1}{y_{22}} & Z_{22} &= -\left[\frac{1}{y_{11}} + \frac{1}{y_{22}} + \frac{y_{12}}{y_{11} y_{22}} \right] \\
 \Delta_Z &= \begin{vmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{vmatrix} = \frac{1}{\Delta_{\beta}} = -\frac{D}{y_{11} y_{22}}
 \end{aligned} \right\}$$

μ = amplification factor.

graphically, showing both phase and magnitude, and this paper is referred to for details.⁴

Tables I, II and III, in conjunction with Figs. 8, 13 and 15, allow us to derive the equivalent networks of Figs. 19, 20 and 21.

We must now undertake a discussion of the results given in the tables as well as of the networks which were derived from them.

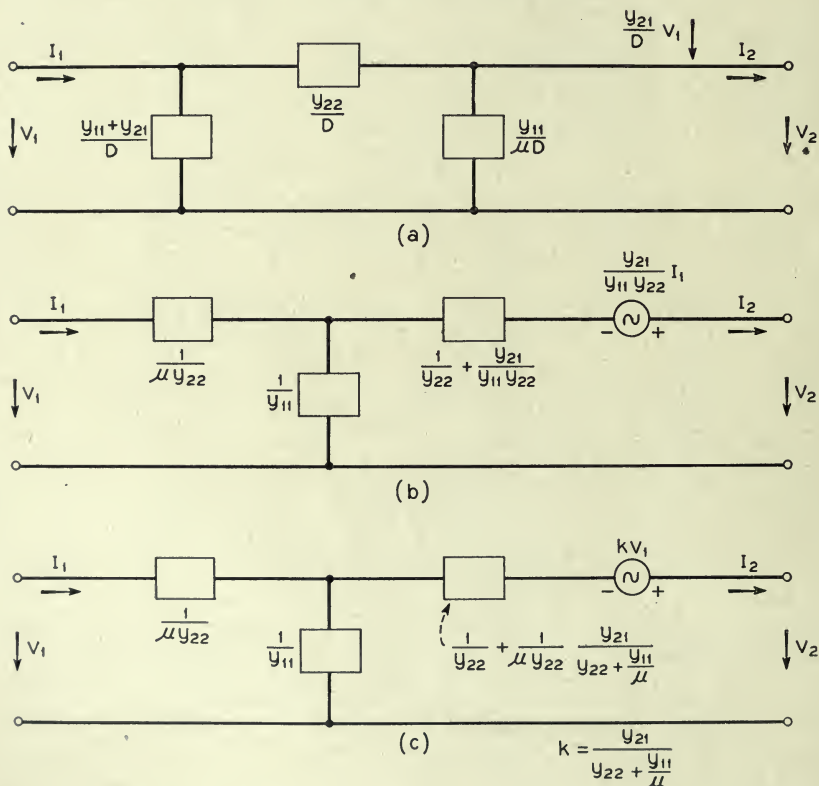


Fig. 19 (a, b, c)—Three forms of equivalent circuits of the grounded cathode triode valid at all frequencies.

Initially, it is well to emphasize again that the four-pole parameters and the corresponding networks are different but equivalent ways through which the triode signal behavior becomes completely specified for all conditions of space charge and for all frequencies.

Secondly, there are certain general relations which should be noted. We observe, first, that the determinants Δ_β and Δ_Z are invariants for the different modes of triode operation, and with exception of phase reversals this

⁴ Loc. cit.

is also the case for the effective transadmittance $\beta_{12} + \beta_{21}$ as well as for the effective transimpedance $Z_{12} + Z_{21}$. On the other hand, the quantity $\frac{Z_{12} + Z_{21}}{Z_{11}}$, which appears in the network of Fig. 15 and which represents the driving force per unit input voltage, is invariant (except for reversal in phase) only for grounded grid and grounded cathode operation. Several

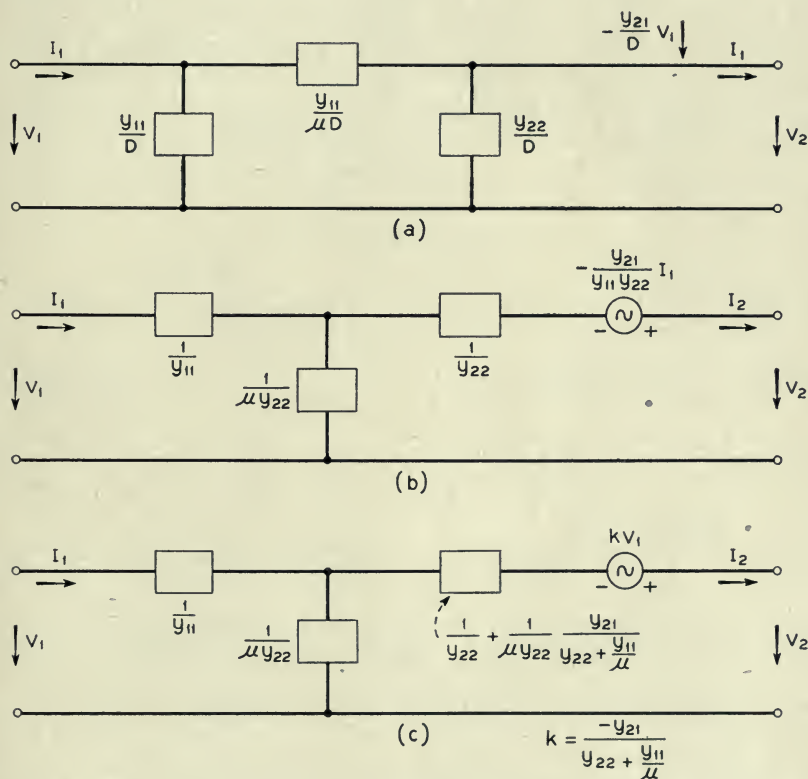


Fig. 20 (a, b, c)—Three forms of equivalent circuits of the grounded grid triode valid at all frequencies.

admittance and impedance relations may also be pointed out. For example, the input short-circuit driving-point admittance β_{11} is equal for grounded cathode and grounded plate operation and the same is true for the output short-circuit driving-point admittances β_{22} for grounded cathode and grounded grid operation. Moreover, it is also seen that the input short-circuit driving-point admittance β_{11} for grounded grid operation is equal to the output driving-point admittance β_{22} for grounded plate operation. A

similar set of reciprocal relations between the open-circuit driving-point impedances is also present.

In regard to the networks it may first be observed that, since they were derived from parameters upon which no restrictions had been placed on either frequency or space charge, they are also generally valid. In passive circuit theory one is accustomed to the use of only the three basic elements

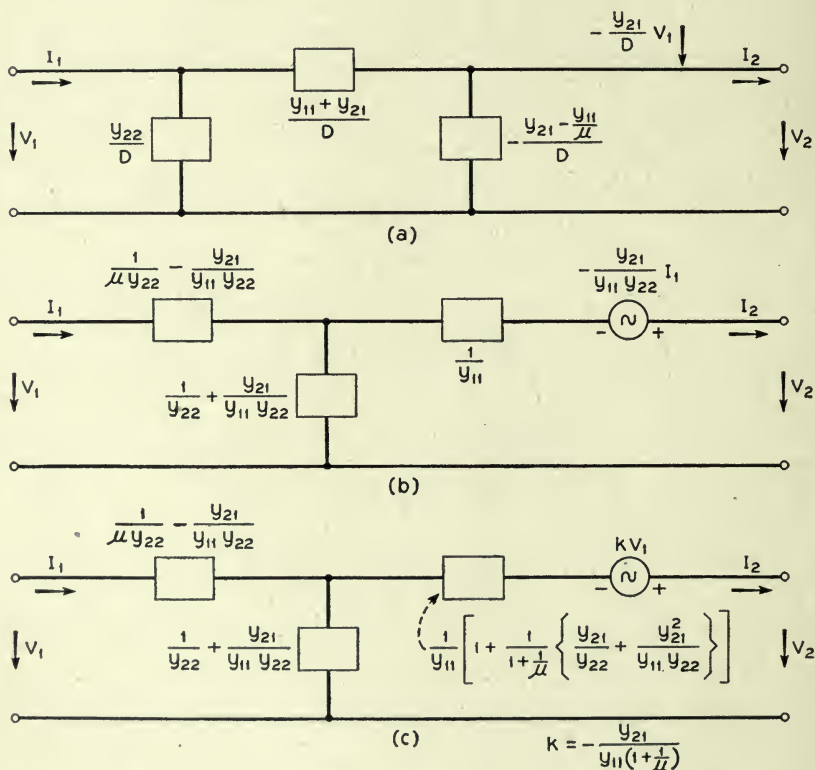


Fig. 21 (a, b, c)—Three forms of equivalent circuits of the grounded plate triode valid at all frequencies.

of resistance, inductance and capacitance. For the networks now under consideration other quantities also need to be included. However, as normally operated there is usually complete space-charge in the cathode-grid region and a very small amount of space charge in the grid-plate region. Under such circumstances the admittance y_{22} is a simple capacitance and the amplification factor μ is a real number. The admittances y_{11} and y_{21} , on the other hand, do not allow accepted circuit interpretation to be made, except in the range of moderately low frequencies when electron transit

time is taken into account only to a first order of approximation. It is believed that, in general, these admittances should be considered complete by themselves as new admittance elements, and, as already remarked, their values in magnitude and phase may be found in the paper on vacuum tube networks to which repeated reference has been made.

As a further property of the networks, consider the expressions for the β 's in Tables I, II and III. It is observed that each β_{ij} of a set of β 's contains a common term, suggesting that the networks might be broken up into at least two elementary constituents. The same observation applies to each of the three sets of Z 's. In Table I, for example, it is seen that this constant term is represented by y_{22}/D . The network of Fig. 19a can thus be thought of as arising from the superposition of two networks, one of which is determined by the parameters

$$\left. \begin{aligned} \beta'_{11} &= \frac{y_{11} + y_{21}}{D} & \beta'_{12} &= 0 \\ \beta'_{21} &= \frac{y_{21}}{D} & \beta'_{22} &= -\frac{y_{11}}{\mu D} \end{aligned} \right\} \quad (25)$$

and the other by the parameters

$$\left. \begin{aligned} \beta''_{11} &= \frac{y_{22}}{D} & \beta''_{12} &= -\frac{y_{22}}{D} \\ \beta''_{21} &= \frac{y_{22}}{D} & \beta''_{22} &= -\frac{y_{22}}{D} \end{aligned} \right\} \quad (26)$$

The admittance coefficients given by (25) correspond to the perfectly unilateral active network shown in Fig. 22a, while the admittance coefficients in (26) correspond to the "passive" network in Fig. 22b. The two elementary constituents thus take the general forms of these networks. It should be noticed that these two network constituents are unrelated to the fact that the total current entering the complete network is the sum of conduction and displacement current.

Corresponding to the Z 's other elementary constituents are obtained, with general forms as shown on Figs. 23a and 23b.

These elementary constituents are merely reflections of certain mathematical identities. For example, the networks in Fig. 22 depend upon the matrix identity

$$\left\| \begin{array}{cc} \beta_{11}\beta_{12} \\ \beta_{21}\beta_{22} \end{array} \right\| \equiv \left\| \begin{array}{cc} \beta_{11} + \beta_{12} & 0 \\ \beta_{21} + \beta_{12} & \beta_{22} - \beta_{12} \end{array} \right\| + \left\| \begin{array}{cc} -\beta_{12}\beta_{12} \\ -\beta_{12}\beta_{12} \end{array} \right\| \quad (27)$$

while those in Fig. 23 depend upon a corresponding identity. The identity (27) expresses the fact that the general " β -network" can be considered as arising from the parallel connections of the two networks in Fig. 24.

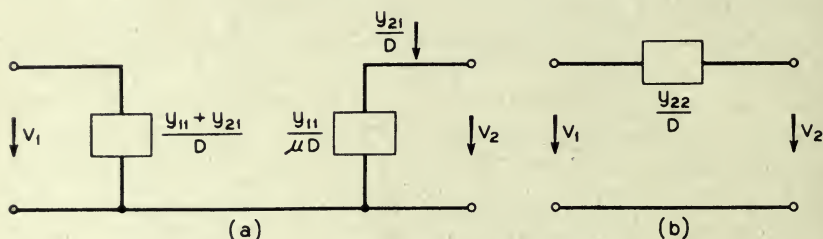


Fig. 22 (a, b)—Elementary constituents of Fig. 19a.

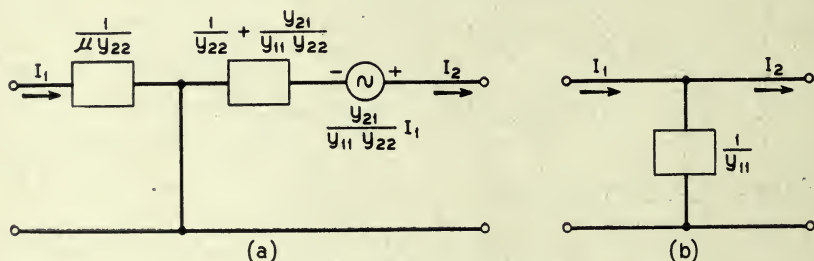


Fig. 23 (a, b)—Elementary constituents of Fig. 19b.

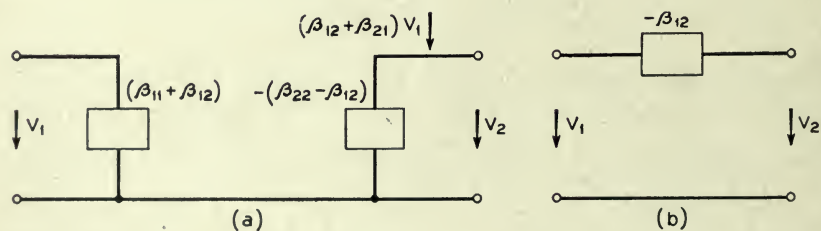


Fig. 24 (a, b)—Elementary constituents of Fig. 8.

There is at present a tendency towards grid designs of very fine mesh. Such a grid design results in a very large value of the amplification factor and, for many purposes, sufficient accuracy may be obtained by disregarding terms containing $\frac{1}{\mu}$ as a factor. Under these conditions the network for grounded cathode operation reduces to an L-network, that for grounded grid operation to a unilateral network transmitting in the direction from grid to plate only, while the cathode follower network remains essentially unchanged.

TABLE IV

FOUR-POLE PARAMETERS FOR GROUNDED CATHODE TRIODE IN THE RANGE OF MODERATELY LOW FREQUENCIES

$$\beta_{11} = \frac{1}{5} \frac{(\omega C_1)^2}{g_0} \frac{F}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2} + i\omega \frac{\frac{4}{3} C_1 f + C_2}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

$$\beta_{12} = - \left[- \frac{(\omega C_1)^2}{g_0} \frac{1}{5\mu} \frac{F}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2} + i\omega \frac{C_2}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

$$\beta_{22} = - \left[\frac{g_0}{\mu} \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} + i\omega \frac{C_2 + \frac{6}{10\mu} C_1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

$$\beta_{12} + \beta_{21} = - \frac{g_0}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \left[1 - i \frac{1}{3} \theta_1 \left(1 - \frac{3}{11\mu} \frac{\frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right) - i \frac{1}{3} \theta_2 k \right]$$

TABLE V

FOUR-POLE PARAMETERS FOR GROUNDED GRID TRIODE IN THE RANGE OF MODERATELY LOW FREQUENCIES

$$\beta_{11} = g_0 \frac{1 + \frac{1}{\mu}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} + i\omega \frac{6}{10} C_1 \frac{1 + \frac{1}{\mu}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \left[1 + \frac{\frac{1}{\mu} \frac{x_2}{x_1}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

$$\beta_{12} = - \frac{g_0}{\mu} \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} - i\omega \frac{6}{10} \frac{C_1}{\mu}$$

$$\beta_{22} = - \left[\frac{g_0}{\mu} \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} + i\omega \left(\frac{C_2 + \frac{6}{10\mu} C_1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right) \right]$$

$$\beta_{12} + \beta_{21} = \frac{g_0}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \left[1 - i \frac{11}{30} \theta_1 \left(1 - \frac{3}{11\mu} \frac{\frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right) - i \frac{1}{3} \theta_2 k \right]$$

In the foregoing, the behavior of an active four-pole has been described either in terms of four-pole parameters or in terms of elements of an equivalent circuit. The particular four-pole parameters, which are of customary use in communication engineering, are the so-called image parameters, but they have usually been used only in connection with passive four-poles. They may, however, also be used in the more general vacuum tube four-pole now under discussion, but whether their employment would be of practical

TABLE VI
FOUR-POLE PARAMETERS FOR GROUNDED PLATE TRIODE IN THE RANGE OF
MODERATELY LOW FREQUENCIES

$$\beta_{11} = \frac{1}{5} \frac{(\omega C_1)^2}{g_0} \frac{F}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2} + i\omega \frac{\frac{4}{3} C_1 f + C_2}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

$$\beta_{12} = - \left[\frac{1}{5} \frac{(\omega C_1)^2}{g_0} \frac{F \left(1 + \frac{1}{\mu}\right)}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2} + i\omega \frac{\frac{4}{3} C_1 f}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

$$\beta_{22} = - \left[g_0 \frac{1 + \frac{1}{\mu}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} + i\omega \frac{6}{10} C_1 \frac{1 + \frac{1}{\mu}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

$$\cdot \left[1 + \frac{1}{3} \frac{\frac{1}{\mu} \frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right]$$

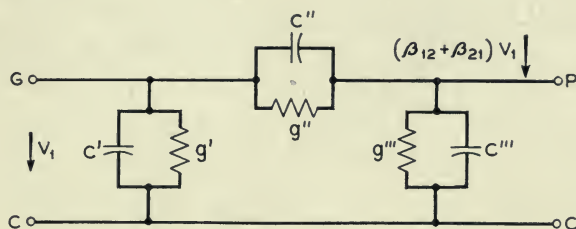
$$\beta_{12} + \beta_{21} = \frac{g_0}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \left[1 - i \frac{11}{30} \theta_1 \left(1 - \frac{3}{11\mu} \frac{\frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \right) - i \frac{1}{3} \theta_2 k \right]$$

value is a matter which engineering experience will decide. In any case their use would be limited to vacuum tubes in which appreciable interaction between input and output terminals is present. From a practical standpoint this means that their usefulness would be mainly found in connection with triodes.

TRIODE NETWORKS AT MODERATELY LOW FREQUENCIES

The networks discussed in the preceding section were of general validity in respect to both operating conditions and frequency and it was mentioned

that the efforts of trying to interpret the "passive" parts of the networks in the form of lumped passive circuit elements had, in general, met with not too much success. In this section attention will be given to the range of moderately low frequencies where usual circuit interpretation is possible. The operating conditions are assumed to be the usual ones with complete



$$g' = \frac{1}{5} \frac{(\omega C_1)^2}{g_0} F \frac{1 + \frac{1}{\mu}}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right) \right]^2}$$

$$C' = \frac{4}{3} C_1 \frac{f}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

$$g'' = - \frac{(\omega C_1)^2}{5g_0} \frac{1}{\mu} \frac{F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

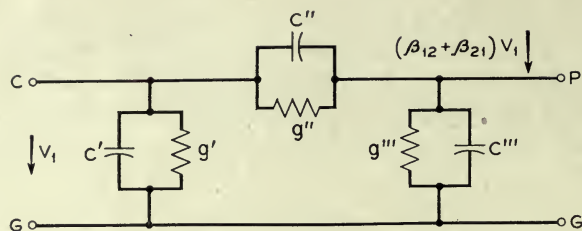
$$C'' = C_2 \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

$$g''' = \frac{g_0}{\mu} \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

Fig. 25—Equivalent circuit of grounded cathode triode at moderately high frequencies.

space charge in the cathode-grid region, and negligible space charge in the grid-plate region. Also it will be assumed that the grid is at negative d-c potential with respect to the cathode.

The first step is to expand the β coefficients in series with transit angles retained only to first or possibly to second orders. As the detailed computations are lengthy only the final result will be given. These are presented in Tables IV, V and VI.



$$g' = g_0 \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

$$C' = \frac{6}{10} C_1 \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)} \left[1 + \frac{1}{3\mu} \frac{\frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)} \right]$$

$$g'' = \frac{1}{\mu} g'$$

$$C'' = \frac{1}{\mu} C'$$

$$g''' = -\frac{(\omega C_1)^2}{5g_0} \frac{1}{\mu} \frac{F}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right) \right]^2}$$

$$C''' = C_2 \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f \right)}$$

Fig. 26—Equivalent circuit of grounded grid triode at moderately high frequencies.

In these tables the symbols have the following meanings:

g_0 = static conductance of the diode formed by the cathode and equivalent grid-plane.

μ = low frequency amplification factor.

x_1 = cathode-grid distance in cm.

x_2 = grid-anode distance in cm.

C_1 = cold capacitance between cathode and equivalent grid plane

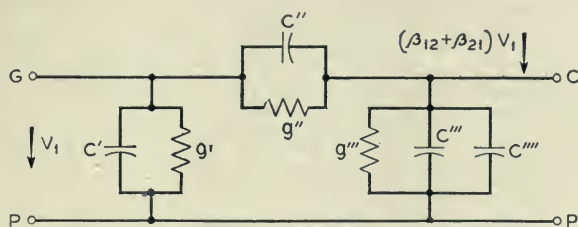
C_2 = cold capacitance between anode and equivalent grid plane.

θ_1 = electron transit angle between cathode and equivalent grid plane.

It is most simply calculated from

$$\theta_1 = 2 \frac{\omega C_1}{g_0}$$

θ_2 = electron transit angle between anode and equivalent grid plane.



$$g' = -\frac{(\omega C_1)^2}{5g_0} \frac{1}{\mu} \frac{F}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2}$$

$$C' = C_2 \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

$$g'' = \frac{(\omega C_1)^2}{5g_0} F \frac{1 + \frac{1}{\mu}}{\left[1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)\right]^2}$$

$$C'' = \frac{4}{3} C_1 \frac{f}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

$$g''' = g_0 \frac{1 + \frac{1}{\mu}}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

$$C''' = \frac{6}{10} \frac{C_1}{\mu} \frac{1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)} \left[1 + \frac{1}{3} \frac{\frac{x_2}{x_1} F}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}\right]$$

$$C'''' = -\frac{11}{15} C_1 \frac{1 + \frac{10}{11} \frac{\theta_2}{\theta_1} k_1}{1 + \frac{1}{\mu} \left(1 + \frac{4}{3} \frac{x_2}{x_1} f\right)}$$

Fig. 27—Equivalent circuit of grounded plate triode at moderately high frequencies.

It may be calculated from

$$\theta_2 = \frac{\omega 2x_2}{\sqrt{2\eta}(\sqrt{V_{D1}} + \sqrt{V_{D2}})}$$

$$\text{where } \eta = 10^7 \frac{e}{m} = 1.76 \times 10^{15}$$

V_{D1} = equivalent d-c grid potential

V_{D2} = anode d-c potential

$$F = 1 + \frac{22}{9} \frac{\theta_2}{\theta_1} \frac{\sqrt{V_{D1}} + 2\sqrt{V_{D2}}}{\sqrt{V_{D1}} + \sqrt{V_{D2}}} + \frac{5}{3} \left(\frac{\theta_2}{\theta_1} \right)^2 \frac{\sqrt{V_{D1}} + 3\sqrt{V_{D2}}}{\sqrt{V_{D1}} + \sqrt{V_{D2}}}$$

$$f = 1 + \frac{1}{2} \frac{\theta_2}{\theta_1} \frac{\sqrt{V_{D1}} + 2\sqrt{V_{D2}}}{\sqrt{V_{D1}} + \sqrt{V_{D2}}}$$

$$k = \frac{\sqrt{V_{D1}} + 2\sqrt{V_{D2}}}{\sqrt{V_{D1}} + \sqrt{V_{D2}}}$$

With the aid of these tables and Fig. 8 the equivalent circuits on Figs. 25, 26 and 27 are obtained. The networks are all of the resistance-capacity type. It may be noted that, in some of the branches, negative conductance or negative capacitance appears. However, as seen from the external tube terminals they are swamped by corresponding positive elements.

The viewpoints presented in this paper have been used by the writer over a number of years. They have been given experimental application by Mr. J. A. Morton, who is principally responsible for their introduction and use in the studies in these Laboratories of electron tubes in the microwave regions.

With this, our investigation comes to a close. Much has been omitted, particularly in the field of applications, but it is nevertheless hoped the fundamental approach, as well as the networks given, may prove to be useful in practical applications. The questions of noise and of optimum noise figure design have also been left out of consideration. Mr. J. A. Morton and the writer plan to discuss these problems in a forthcoming paper.

The writer is pleased to acknowledge his indebtedness to Messrs. R. K. Potter, J. A. Morton, and R. M. Ryder, who have encouraged this work and urged its publication; and to Mr. W. E. Kirkpatrick for constructively critical scrutiny of the original technical memorandum.

A Mathematical Theory of Communication

By C. E. SHANNON

(Concluded from July 1948 issue)

PART III: MATHEMATICAL PRELIMINARIES

In this final installment of the paper we consider the case where the signals or the messages or both are continuously variable, in contrast with the discrete nature assumed until now. To a considerable extent the continuous case can be obtained through a limiting process from the discrete case by dividing the continuum of messages and signals into a large but finite number of small regions and calculating the various parameters involved on a discrete basis. As the size of the regions is decreased these parameters in general approach as limits the proper values for the continuous case. There are, however, a few new effects that appear and also a general change of emphasis in the direction of specialization of the general results to particular cases.

We will not attempt, in the continuous case, to obtain our results with the greatest generality, or with the extreme rigor of pure mathematics, since this would involve a great deal of abstract measure theory and would obscure the main thread of the analysis. A preliminary study, however, indicates that the theory can be formulated in a completely axiomatic and rigorous manner which includes both the continuous and discrete cases and many others. The occasional liberties taken with limiting processes in the present analysis can be justified in all cases of practical interest.

18. SETS AND ENSEMBLES OF FUNCTIONS

We shall have to deal in the continuous case with sets of functions and ensembles of functions. A set of functions, as the name implies, is merely a class or collection of functions, generally of one variable, time. It can be specified by giving an explicit representation of the various functions in the set, or implicitly by giving a property which functions in the set possess and others do not. Some examples are:

1. The set of functions:

$$f_{\theta}(t) = \sin(t + \theta).$$

Each particular value of θ determines a particular function in the set.

2. The set of all functions of time containing no frequencies over W cycles per second.
3. The set of all functions limited in band to W and in amplitude to A .
4. The set of all English speech signals as functions of time.

An *ensemble* of functions is a set of functions together with a probability measure whereby we may determine the probability of a function in the set having certain properties.¹ For example with the set,

$$f_{\theta}(t) = \sin(t + \theta),$$

we may give a probability distribution for θ , $P(\theta)$. The set then becomes an ensemble.

Some further examples of ensembles of functions are:

1. A finite set of functions $f_k(t)$ ($k = 1, 2, \dots, n$) with the probability of f_k being p_k .
2. A finite dimensional family of functions

$$f(\alpha_1, \alpha_2, \dots, \alpha_n; t)$$

with a probability distribution for the parameters α_i :

$$p(\alpha_1, \dots, \alpha_n)$$

For example we could consider the ensemble defined by

$$f(a_1, \dots, a_n, \theta_1, \dots, \theta_n; t) = \sum_{n=1}^n a_n \sin n(\omega t + \theta_n)$$

with the amplitudes a_i distributed normally and independently, and the phases θ_i distributed uniformly (from 0 to 2π) and independently.

3. The ensemble

$$f(a_i, t) = \sum_{n=-\infty}^{+\infty} a_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}$$

with the a_i normal and independent all with the same standard deviation $\sqrt{N}$. This is a representation of "white" noise, band-limited to the band from 0 to W cycles per second and with average power N .²

¹ In mathematical terminology the functions belong to a measure space whose total measure is unity.

² This representation can be used as a definition of band limited white noise. It has certain advantages in that it involves fewer limiting operations than do definitions that have been used in the past. The name "white noise," already firmly entrenched in the literature, is perhaps somewhat unfortunate. In optics white light means either any continuous spectrum as contrasted with a point spectrum, or a spectrum which is flat with *wavelength* (which is not the same as a spectrum flat with frequency).

4. Let points be distributed on the t axis according to a Poisson distribution. At each selected point the function $f(t)$ is placed and the different functions added, giving the ensemble

$$\sum_{k=-\infty}^{\infty} f(t + t_k)$$

where the t_k are the points of the Poisson distribution. This ensemble can be considered as a type of impulse or shot noise where all the impulses are identical.

5. The set of English speech functions with the probability measure given by the frequency of occurrence in ordinary use.

An ensemble of functions $f_{\alpha}(t)$ is *stationary* if the same ensemble results when all functions are shifted any fixed amount in time. The ensemble

$$f_{\theta}(t) = \sin(t + \theta)$$

is stationary if θ distributed uniformly from 0 to 2π . If we shift each function by t_1 we obtain

$$\begin{aligned} f_{\theta}(t + t_1) &= \sin(t + t_1 + \theta) \\ &= \sin(t + \varphi) \end{aligned}$$

with φ distributed uniformly from 0 to 2π . Each function has changed but the ensemble as a whole is invariant under the translation. The other examples given above are also stationary.

An ensemble is *ergodic* if it is stationary, and there is no subset of the functions in the set with a probability different from 0 and 1 which is stationary. The ensemble

$$\sin(t + \theta)$$

is ergodic. No subset of these functions of probability $\neq 0, 1$ is transformed into itself under all time translations. On the other hand the ensemble

$$a \sin(t + \theta)$$

with a distributed normally and θ uniform is stationary but not ergodic. The subset of these functions with a between 0 and 1 for example is stationary.

Of the examples given, 3 and 4 are ergodic, and 5 may perhaps be considered so. If an ensemble is ergodic we may say roughly that each function in the set is typical of the ensemble. More precisely it is known that with an ergodic ensemble an average of any statistic over the ensemble is equal (with probability 1) to an average over all the time translations of a

particular function in the set.³ Roughly speaking, each function can be expected, as time progresses, to go through, with the proper frequency, all the convolutions of any of the functions in the set.

Just as we may perform various operations on numbers or functions to obtain new numbers or functions, we can perform operations on ensembles to obtain new ensembles. Suppose, for example, we have an ensemble of functions $f_\alpha(t)$ and an operator T which gives for each function $f_\alpha(t)$ a result $g_\alpha(t)$:

$$g_\alpha(t) = Tf_\alpha(t)$$

Probability measure is defined for the set $g_\alpha(t)$ by means of that for the set $f_\alpha(t)$. The probability of a certain subset of the $g_\alpha(t)$ functions is equal to that of the subset of the $f_\alpha(t)$ functions which produce members of the given subset of g functions under the operation T . Physically this corresponds to passing the ensemble through some device, for example, a filter, a rectifier or a modulator. The output functions of the device form the ensemble $g_\alpha(t)$.

A device or operator T will be called invariant if shifting the input merely shifts the output, i.e., if

$$g_\alpha(t) = Tf_\alpha(t)$$

implies

$$g_\alpha(t + t_1) = Tf_\alpha(t + t_1)$$

for all $f_\alpha(t)$ and all t_1 . It is easily shown (see appendix 1) that if T is invariant and the input ensemble is stationary then the output ensemble is stationary. Likewise if the input is ergodic the output will also be ergodic.

A filter or a rectifier is invariant under all time translations. The operation of modulation is not since the carrier phase gives a certain time structure. However, modulation is invariant under all translations which are multiples of the period of the carrier.

Wiener has pointed out the intimate relation between the invariance of physical devices under time translations and Fourier theory.⁴ He has

³ This is the famous ergodic theorem or rather one aspect of this theorem which was proved is somewhat different formulations by Birkhoff, von Neumann, and Koopman, and subsequently generalized by Wiener, Hopf, Hurewicz and others. The literature on ergodic theory is quite extensive and the reader is referred to the papers of these writers for precise and general formulations; e.g., E. Hopf "Ergodentheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Vol. 5, "On Causality Statistics and Probability" *Journal of Mathematics and Physics*, Vol. XIII, No. 1, 1934; N. Wiener "The Ergodic Theorem" *Duke Mathematical Journal*, Vol. 5, 1939.

⁴ Communication theory is heavily indebted to Wiener for much of its basic philosophy and theory. His classic NDRC report "The Interpolation, Extrapolation, and Smoothing of Stationary Time Series," to appear soon in book form, contains the first clear-cut formulation of communication theory as a statistical problem, the study of operations

shown, in fact, that if a device is linear as well as invariant Fourier analysis is then the appropriate mathematical tool for dealing with the problem.

An ensemble of functions is the appropriate mathematical representation of the messages produced by a continuous source (for example speech), of the signals produced by a transmitter, and of the perturbing noise. Communication theory is properly concerned, as has been emphasized by Wiener, not with operations on particular functions, but with operations on ensembles of functions. A communication system is designed not for a particular speech function and still less for a sine wave, but for the ensemble of speech functions.

19. BAND LIMITED ENSEMBLES OF FUNCTIONS

If a function of time $f(t)$ is limited to the band from 0 to W cycles per second it is completely determined by giving its ordinates at a series of discrete points spaced $\frac{1}{2W}$ seconds apart in the manner indicated by the following result.⁵

Theorem 13: Let $f(t)$ contain no frequencies over W .

Then

$$f(t) = \sum_{-\infty}^{\infty} X_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}$$

where

$$X_n = f\left(\frac{n}{2W}\right).$$

In this expansion $f(t)$ is represented as a sum of orthogonal functions. The coefficients X_n of the various terms can be considered as coordinates in an infinite dimensional "function space." In this space each function corresponds to precisely one point and each point to one function.

A function can be considered to be substantially limited to a time T if all the ordinates X_n outside this interval of time are zero. In this case all but $2TW$ of the coordinates will be zero. Thus functions limited to a band W and duration T correspond to points in a space of $2TW$ dimensions.

A subset of the functions of band W and duration T corresponds to a region in this space. For example, the functions whose total energy is less

on time series. This work, although chiefly concerned with the linear prediction and filtering problem, is an important collateral reference in connection with the present paper. We may also refer here to Wiener's forthcoming book "Cybernetics" dealing with the general problems of communication and control.

⁵ For a proof of this theorem and further discussion see the author's paper "Communication in the Presence of Noise" to be published in the *Proceedings of the Institute of Radio Engineers*.

than or equal to E correspond to points in a $2TW$ dimensional sphere with radius $r = \sqrt{2WE}$.

An *ensemble* of functions of limited duration and band will be represented by a probability distribution $p(x_1 \cdots x_n)$ in the corresponding n dimensional space. If the ensemble is not limited in time we can consider the $2TW$ coordinates in a given interval T to represent substantially the part of the function in the interval T and the probability distribution $p(x_1, \cdots, x_n)$ to give the statistical structure of the ensemble for intervals of that duration.

20. ENTROPY OF A CONTINUOUS DISTRIBUTION

The entropy of a discrete set of probabilities $p_1, \cdots p_n$ has been defined as:

$$H = -\sum p_i \log p_i.$$

In an analogous manner we define the entropy of a continuous distribution with the density distribution function $p(x)$ by:

$$H = -\int_{-\infty}^{\infty} p(x) \log p(x) dx$$

With an n dimensional distribution $p(x_1, \cdots, x_n)$ we have

$$H = -\int \cdots \int p(x_1 \cdots x_n) \log p(x_1, \cdots, x_n) dx_1 \cdots dx_n.$$

If we have two arguments x and y (which may themselves be multi-dimensional) the joint and conditional entropies of $p(x, y)$ are given by

$$H(x, y) = -\iint p(x, y) \log p(x, y) dx dy$$

and

$$H_x(y) = -\iint p(x, y) \log \frac{p(x, y)}{p(x)} dx dy$$

$$H_y(x) = -\iint p(x, y) \log \frac{p(x, y)}{p(y)} dx dy$$

where

$$p(x) = \int p(x, y) dy$$

$$p(y) = \int p(x, y) dx.$$

The entropy of continuous distributions have most (but not all) of the properties of the discrete case. In particular we have the following:

1. If x is limited to a certain volume v in its space, then $H(x)$ is a maximum and equal to $\log v$ when $p(x)$ is constant $\left(\frac{1}{v}\right)$ in the volume.
2. With any two variables x, y we have

$$H(x, y) \leq H(x) + H(y)$$

with equality if (and only if) x and y are independent, i.e., $p(x, y) = p(x)p(y)$ (apart possibly from a set of points of probability zero).

3. Consider a generalized averaging operation of the following type:

$$p'(y) = \int a(x, y)p(x) dx$$

with

$$\int a(x, y) dx = \int a(x, y) dy = 1, \quad a(x, y) \geq 0.$$

Then the entropy of the averaged distribution $p'(y)$ is equal to or greater than that of the original distribution $p(x)$.

4. We have

$$H(x, y) = H(x) + H_x(y) = H(y) + H_y(x)$$

and

$$H_x(y) \leq H(y).$$

5. Let $p(x)$ be a one-dimensional distribution. The form of $p(x)$ giving a maximum entropy subject to the condition that the standard deviation of x be fixed at σ is gaussian. To show this we must maximize

$$H(x) = - \int p(x) \log p(x) dx$$

with

$$\sigma^2 = \int p(x)x^2 dx \quad \text{and} \quad 1 = \int p(x) dx$$

as constraints. This requires, by the calculus of variations, maximizing

$$\int [-p(x) \log p(x) + \lambda p(x)x^2 + \mu p(x)] dx.$$

The condition for this is

$$-1 - \log p(x) + \lambda x^2 + \mu = 0$$

and consequently (adjusting the constants to satisfy the constraints)

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)}.$$

Similarly in n dimensions, suppose the second order moments of $p(x_1, \dots, x_n)$ are fixed at A_{ij} :

$$A_{ij} = \int \cdots \int x_i x_j p(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Then the maximum entropy occurs (by a similar calculation) when $p(x_1, \dots, x_n)$ is the n dimensional gaussian distribution with the second order moments A_{ij} .

6. The entropy of a one-dimensional gaussian distribution whose standard deviation is σ is given by

$$H(x) = \log \sqrt{2\pi e} \sigma.$$

This is calculated as follows:

$$\begin{aligned} p(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-(x^2/2\sigma^2)} \\ -\log p(x) &= \log \sqrt{2\pi}\sigma + \frac{x^2}{2\sigma^2} \\ H(x) &= -\int p(x) \log p(x) dx \\ &= \int p(x) \log \sqrt{2\pi}\sigma dx + \int p(x) \frac{x^2}{2\sigma^2} dx \\ &= \log \sqrt{2\pi}\sigma + \frac{\sigma^2}{2\sigma^2} \\ &= \log \sqrt{2\pi}\sigma + \log \sqrt{e} \\ &= \log \sqrt{2\pi e} \sigma. \end{aligned}$$

Similarly the n dimensional gaussian distribution with associated quadratic form a_{ij} is given by

$$p(x_1, \dots, x_n) = \frac{|a_{ij}|^{\frac{1}{2}}}{(2\pi)^{n/2}} \exp(-\frac{1}{2} \sum a_{ij} X_i X_j)$$

and the entropy can be calculated as

$$H = \log (2\pi e)^{n/2} |a_{ij}|^{\frac{1}{2}}$$

where $|a_{ij}|$ is the determinant whose elements are a_{ij} .

7. If x is limited to a half line ($p(x) = 0$ for $x \leq 0$) and the first moment of x is fixed at a :

$$a = \int_0^\infty p(x)x dx,$$

then the maximum entropy occurs when

$$p(x) = \frac{1}{a} e^{-(x/a)}$$

and is equal to $\log ea$.

8. There is one important difference between the continuous and discrete entropies. In the discrete case the entropy measures in an *absolute* way the randomness of the chance variable. In the continuous case the measurement is *relative to the coordinate system*. If we change coordinates the entropy will in general change. In fact if we change to coordinates $y_1 \cdots y_n$ the new entropy is given by

$$H(y) = \int \cdots \int p(x_1 \cdots x_n) J \left(\frac{x}{y} \right) \log p(x_1 \cdots x_n) J \left(\frac{x}{y} \right) dy_1 \cdots dy_n$$

where $J \left(\frac{x}{y} \right)$ is the Jacobian of the coordinate transformation. On expanding the logarithm and changing variables to $x_1 \cdots x_n$, we obtain:

$$H(y) = H(x) - \int \cdots \int p(x_1, \cdots, x_n) \log J \left(\frac{x}{y} \right) dx_1 \cdots dx_n.$$

Thus the new entropy is the old entropy less the expected logarithm of the Jacobian. In the continuous case the entropy can be considered a measure of randomness *relative to an assumed standard*, namely the coordinate system chosen with each small volume element $dx_1 \cdots dx_n$ given equal weight. When we change the coordinate system the entropy in the new system measures the randomness when equal volume elements $dy_1 \cdots dy_n$ in the new system are given equal weight.

In spite of this dependence on the coordinate system the entropy concept is as important in the continuous case as the discrete case. This is due to the fact that the derived concepts of information rate and channel capacity depend on the *difference* of two entropies and this difference *does not* depend on the coordinate frame, each of the two terms being changed by the same amount.

The entropy of a continuous distribution can be negative. The scale of measurements sets an arbitrary zero corresponding to a uniform distribution over a unit volume. A distribution which is more confined than this has less entropy and will be negative. The rates and capacities will, however, always be non-negative.

9. A particular case of changing coordinates is the linear transformation

$$y_j = \sum_i a_{ij} x_i.$$

In this case the Jacobian is simply the determinant $|a_{ij}|^{-1}$ and

$$H(y) = H(x) + \log |a_{ij}|.$$

In the case of a rotation of coordinates (or any measure preserving transformation) $J = 1$ and $H(y) = H(x)$.

21. ENTROPY OF AN ENSEMBLE OF FUNCTIONS

Consider an ergodic ensemble of functions limited to a certain band of width W cycles per second. Let

$$p(x_1 \cdots x_n)$$

be the density distribution function for amplitudes $x_1 \cdots x_n$ at n successive sample points. We define the entropy of the ensemble per degree of freedom by

$$H' = -\lim_{n \rightarrow \infty} \frac{1}{n} \int \cdots \int p(x_1 \cdots x_n) \log p(x_1, \cdots, x_n) dx_1 \cdots dx_n.$$

We may also define an entropy H per second by dividing, not by n , but by the time T in seconds for n samples. Since $n = 2TW$, $H' = 2WH$.

With white thermal noise p is gaussian and we have

$$H' = \log \sqrt{2\pi eN},$$

$$H = W \log 2\pi eN.$$

For a given average power N , white noise has the maximum possible entropy. This follows from the maximizing properties of the Gaussian distribution noted above.

The entropy for a continuous stochastic process has many properties analogous to that for discrete processes. In the discrete case the entropy was related to the logarithm of the *probability* of long sequences, and to the *number* of reasonably probable sequences of long length. In the continuous case it is related in a similar fashion to the logarithm of the *probability density* for a long series of samples, and the *volume* of reasonably high probability in the function space.

More precisely, if we assume $p(x_1 \cdots x_n)$ continuous in all the x_i for all n , then for sufficiently large n

$$\left| \frac{\log p}{n} - H' \right| < \epsilon$$

for all choices of $(x_1, \cdots, x_n)$ apart from a set whose total probability is less than δ , with δ and ϵ arbitrarily small. This follows from the ergodic property if we divide the space into a large number of small cells.

The relation of H to volume can be stated as follows: Under the same assumptions consider the n dimensional space corresponding to $p(x_1, \dots, x_n)$. Let $V_n(q)$ be the smallest volume in this space which includes in its interior a total probability q . Then

$$\lim_{n \rightarrow \infty} \frac{\log V_n(q)}{n} = H'$$

provided q does not equal 0 or 1.

These results show that for large n there is a rather well-defined volume (at least in the logarithmic sense) of high probability, and that within this volume the probability density is relatively uniform (again in the logarithmic sense).

In the white noise case the distribution function is given by

$$p(x_1 \cdots x_n) = \frac{1}{(2\pi N)^{n/2}} \exp - \frac{1}{2N} \sum x_i^2.$$

Since this depends only on $\sum x_i^2$ the surfaces of equal probability density are spheres and the entire distribution has spherical symmetry. The region of high probability is a sphere of radius $\sqrt{nN}$. As $n \rightarrow \infty$ the probability of being outside a sphere of radius $\sqrt{n(N + \epsilon)}$ approaches zero and $\frac{1}{n}$ times the logarithm of the volume of the sphere approaches $\log \sqrt{2\pi e N}$.

In the continuous case it is convenient to work not with the entropy H of an ensemble but with a derived quantity which we will call the entropy power. This is defined as the power in a white noise limited to the same band as the original ensemble and having the same entropy. In other words if H' is the entropy of an ensemble its entropy power is

$$N_1 = \frac{1}{2\pi e} \exp 2H'.$$

In the geometrical picture this amounts to measuring the high probability volume by the squared radius of a sphere having the same volume. Since white noise has the maximum entropy for a given power, the entropy power of any noise is less than or equal to its actual power.

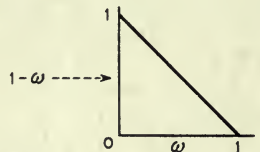
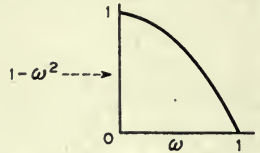
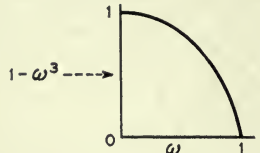
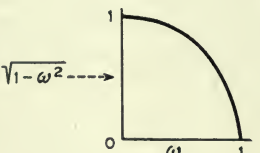
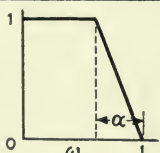
21. ENTROPY LOSS IN LINEAR FILTERS

Theorem 14: If an ensemble having an entropy H_1 per degree of freedom in band W is passed through a filter with characteristic $Y(f)$ the output ensemble has an entropy

$$H_2 = H_1 + \frac{1}{W} \int_W \log |Y(f)|^2 df.$$

The operation of the filter is essentially a linear transformation of coordinates. If we think of the different frequency components as the original coordinate system, the new frequency components are merely the old ones multiplied by factors. The coordinate transformation matrix is thus es-

TABLE I

GAIN	ENTROPY POWER FACTOR	ENTROPY POWER GAIN IN DECIBELS	IMPULSE RESPONSE
	$\frac{1}{e^2}$	-8.68	$\frac{\sin^2 \pi t}{(\pi t)^2}$
	$\left(\frac{2}{e}\right)^4$	-5.32	$2 \left[\frac{\sin t}{t^3} - \frac{\cos t}{t^2} \right]$
	0.384	-4.15	$6 \left[\frac{\cos t - 1}{t^4} - \frac{\cos t}{2t^2} + \frac{\sin t}{t^3} \right]$
	$\left(\frac{2}{e}\right)^2$	-2.66	$\frac{\pi}{2} \frac{J_1(t)}{t}$
	$\frac{1}{e^{2\alpha}}$	-8.68 α	$\frac{1}{\alpha t^2} \left[\cos(1-\alpha)t - \cos t \right]$

entially diagonalized in terms of these coordinates. The Jacobian of the transformation is (for n sine and n cosine components)

$$J = \prod_{i=1}^n |Y(f_i)|^2$$

where the f_i are equally spaced through the band W . This becomes in the limit

$$\exp \frac{1}{W} \int_W \log |Y(f)|^2 df.$$

Since J is constant its average value is this same quantity and applying the theorem on the change of entropy with a change of coordinates, the result follows. We may also phrase it in terms of the entropy power. Thus if the entropy power of the first ensemble is N_1 that of the second is

$$N_1 \exp \frac{1}{W} \int_W \log |Y(f)|^2 df.$$

The final entropy power is the initial entropy power multiplied by the geometric mean gain of the filter. If the gain is measured in db , then the output entropy power will be increased by the arithmetic mean db gain over W .

In Table I the entropy power loss has been calculated (and also expressed in db) for a number of ideal gain characteristics. The impulsive responses of these filters are also given for $W = 2\pi$, with phase assumed to be 0.

The entropy loss for many other cases can be obtained from these results.

For example the entropy power factor $\frac{1}{e^2}$ for the first case also applies to any gain characteristic obtained from $1 - \omega$ by a measure preserving transformation of the ω axis. In particular a linearly increasing gain $G(\omega) = \omega$, or a "saw tooth" characteristic between 0 and 1 have the same entropy loss.

The reciprocal gain has the reciprocal factor. Thus $\frac{1}{\omega}$ has the factor e^2 .

Raising the gain to any power raises the factor to this power.

22. ENTROPY OF THE SUM OF TWO ENSEMBLES

If we have two ensembles of functions $f_\alpha(t)$ and $g_\beta(t)$ we can form a new ensemble by "addition." Suppose the first ensemble has the probability density function $p(x_1, \dots, x_n)$ and the second $q(x_1, \dots, x_n)$. Then the density function for the sum is given by the convolution:

$$r(x_1, \dots, x_n) = \int \dots \int p(y_1, \dots, y_n) \cdot q(x_1 - y_1, \dots, x_n - y_n) dy_1, dy_2, \dots, dy_n.$$

Physically this corresponds to adding the noises or signals represented by the original ensembles of functions.

The following result is derived in Appendix 6.

Theorem 15: Let the average power of two ensembles be N_1 and N_2 and let their entropy powers be $\bar{N}_1$ and $\bar{N}_2$. Then the entropy power of the sum, $\bar{N}_3$, is bounded by

$$\bar{N}_1 + \bar{N}_2 \leq \bar{N}_3 \leq N_1 + N_2.$$

White Gaussian noise has the peculiar property that it can absorb any other noise or signal ensemble which may be added to it with a resultant entropy power approximately equal to the sum of the white noise power and the signal power (measured from the average signal value, which is normally zero), provided the signal power is small, in a certain sense, compared to the noise.

Consider the function space associated with these ensembles having n dimensions. The white noise corresponds to a spherical Gaussian distribution in this space. The signal ensemble corresponds to another probability distribution, not necessarily Gaussian or spherical. Let the second moments of this distribution about its center of gravity be a_{ij} . That is, if $p(x_1, \dots, x_n)$ is the density distribution function

$$a_{ij} = \int \cdots \int p(x_i - \alpha_i)(x_j - \alpha_j) dx_1, \dots, dx_n$$

where the α_i are the coordinates of the center of gravity. Now a_{ij} is a positive definite quadratic form, and we can rotate our coordinate system to align it with the principal directions of this form. a_{ij} is then reduced to diagonal form b_{ii} . We require that each b_{ii} be small compared to N , the squared radius of the spherical distribution.

In this case the convolution of the noise and signal produce a Gaussian distribution whose corresponding quadratic form is

$$N + b_{ii}.$$

The entropy power of this distribution is

$$[\Pi(N + b_{ii})]^{1/n}$$

or approximately

$$\begin{aligned} &= [(N)^n + \Sigma b_{ii}(N)^{n-1}]^{1/n} \\ &\doteq N + \frac{1}{n} \Sigma b_{ii}. \end{aligned}$$

The last term is the signal power, while the first is the noise power.

PART IV: THE CONTINUOUS CHANNEL

23. THE CAPACITY OF A CONTINUOUS CHANNEL

In a continuous channel the input or transmitted signals will be continuous functions of time $f(t)$ belonging to a certain set, and the output or received signals will be perturbed versions of these. We will consider only the case where both transmitted and received signals are limited to a certain band W . They can then be specified, for a time T , by $2TW$ numbers, and their statistical structure by finite dimensional distribution functions. Thus the statistics of the transmitted signal will be determined by

$$P(x_1, \dots, x_n) = P(x)$$

and those of the noise by the conditional probability distribution

$$P_{x_1, \dots, x_n}(y_1, \dots, y_n) = P_x(y).$$

The rate of transmission of information for a continuous channel is defined in a way analogous to that for a discrete channel, namely

$$R = H(x) - H_y(x)$$

where $H(x)$ is the entropy of the input and $H_y(x)$ the equivocation. The channel capacity C is defined as the maximum of R when we vary the input over all possible ensembles. This means that in a finite dimensional approximation we must vary $P(x) = P(x_1, \dots, x_n)$ and maximize

$$- \int P(x) \log P(x) dx + \int \int P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy.$$

This can be written

$$\int \int P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

using the fact that $\int \int P(x, y) \log P(x) dx dy = \int P(x) \log P(x) dx$. The channel capacity is thus expressed

$$C = \lim_{T \rightarrow \infty} \max_{P(x)} \frac{1}{T} \int \int P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy.$$

It is obvious in this form that R and C are independent of the coordinate system since the numerator and denominator in $\log \frac{P(x, y)}{P(x)P(y)}$ will be multiplied by the same factors when x and y are transformed in any one to one way. This integral expression for C is more general than $H(x) - H_y(x)$. Properly interpreted (see Appendix 7) it will always exist while $H(x) - H_y(x)$

may assume an indeterminate form $\infty - \infty$ in some cases. This occurs, for example, if x is limited to a surface of fewer dimensions than n in its n dimensional approximation.

If the logarithmic base used in computing $H(x)$ and $H_y(x)$ is two then C is the maximum number of binary digits that can be sent per second over the channel with arbitrarily small equivocation, just as in the discrete case. This can be seen physically by dividing the space of signals into a large number of small cells, sufficiently small so that the probability density $P_x(y)$ of signal x being perturbed to point y is substantially constant over a cell (either of x or y). If the cells are considered as distinct points the situation is essentially the same as a discrete channel and the proofs used there will apply. But it is clear physically that this quantizing of the volume into individual points cannot in any practical situation alter the final answer significantly, provided the regions are sufficiently small. Thus the capacity will be the limit of the capacities for the discrete subdivisions and this is just the continuous capacity defined above.

On the mathematical side it can be shown first (see Appendix 7) that if u is the message, x is the signal, y is the received signal (perturbed by noise) and v the recovered message then

$$H(x) - H_y(x) \geq H(u) - H_v(u)$$

regardless of what operations are performed on u to obtain x or on y to obtain v . Thus no matter how we encode the binary digits to obtain the signal, or how we decode the received signal to recover the message, the discrete rate for the binary digits does not exceed the channel capacity we have defined. On the other hand, it is possible under very general conditions to find a coding system for transmitting binary digits at the rate C with as small an equivocation or frequency of errors as desired. This is true, for example, if, when we take a finite dimensional approximating space for the signal functions, $P(x, y)$ is continuous in both x and y except at a set of points of probability zero.

An important special case occurs when the noise is added to the signal and is independent of it (in the probability sense). Then $P_x(y)$ is a function only of the difference $n = (y - x)$,

$$P_x(y) = Q(y - x)$$

and we can assign a definite entropy to the noise (independent of the statistics of the signal), namely the entropy of the distribution $Q(n)$. This entropy will be denoted by $H(n)$.

Theorem 16: If the signal and noise are independent and the received signal is the sum of the transmitted signal and the noise then the rate of

transmission is

$$R = H(y) - H(n)$$

i.e., the entropy of the received signal less the entropy of the noise. The channel capacity is

$$C = \text{Max}_{P(x)} H(y) - H(n).$$

We have, since $y = x + n$:

$$H(x, y) = H(x, n).$$

Expanding the left side and using the fact that x and n are independent

$$H(y) + H_y(x) = H(x) + H(n).$$

Hence

$$R = H(x) - H_y(x) = H(y) - H(n).$$

Since $H(n)$ is independent of $P(x)$, maximizing R requires maximizing $H(y)$, the entropy of the received signal. If there are certain constraints on the ensemble of transmitted signals, the entropy of the received signal must be maximized subject to these constraints.

24. CHANNEL CAPACITY WITH AN AVERAGE POWER LIMITATION

A simple application of Theorem 16 is the case where the noise is a white thermal noise and the transmitted signals are limited to a certain average power P . Then the received signals have an average power $P + N$ where N is the average noise power. The maximum entropy for the received signals occurs when they also form a white noise ensemble since this is the greatest possible entropy for a power $P + N$ and can be obtained by a suitable choice of the ensemble of transmitted signals, namely if they form a white noise ensemble of power P . The entropy (per second) of the received ensemble is then

$$H(y) = W \log 2\pi e(P + N),$$

and the noise entropy is

$$H(n) = W \log 2\pi eN.$$

The channel capacity is

$$C = H(y) - H(n) = W \log \frac{P + N}{N}.$$

Summarizing we have the following:

Theorem 17: The capacity of a channel of band W perturbed by white

thermal noise of power N when the average transmitter power is P is given by

$$C = W \log \frac{P + N}{N}.$$

This means of course that by sufficiently involved encoding systems we can transmit binary digits at the rate $W \log_2 \frac{P + N}{N}$ bits per second, with arbitrarily small frequency of errors. It is not possible to transmit at a higher rate by any encoding system without a definite positive frequency of errors.

* To approximate this limiting rate of transmission the transmitted signals must approximate, in statistical properties, a white noise.⁶ A system which approaches the ideal rate may be described as follows: Let $M = 2^s$ samples of white noise be constructed each of duration T . These are assigned binary numbers from 0 to $(M - 1)$. At the transmitter the message sequences are broken up into groups of s and for each group the corresponding noise sample is transmitted as the signal. At the receiver the M samples are known and the actual received signal (perturbed by noise) is compared with each of them. The sample which has the least R.M.S. discrepancy from the received signal is chosen as the transmitted signal and the corresponding binary number reconstructed. This process amounts to choosing the most probable (*a posteriori*) signal. The number M of noise samples used will depend on the tolerable frequency ϵ of errors, but for almost all selections of samples we have

$$\lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log M(\epsilon, T)}{T} = W \log \frac{P + N}{N},$$

so that no matter how small ϵ is chosen, we can, by taking T sufficiently large, transmit as near as we wish to $TW \log \frac{P + N}{N}$ binary digits in the time T .

Formulas similar to $C = W \log \frac{P + N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,⁷ W. G. Tuller,⁸ and H. Sullivan in this connection.

In the case of an arbitrary perturbing noise (not necessarily white thermal noise) it does not appear that the maximizing problem involved in deter-

⁶ This and other properties of the white noise case are discussed from the geometrical point of view in "Communication in the Presence of Noise," loc. cit.

⁷ "Cybernetics," loc. cit.

⁸ Sc. D. thesis, Department of Electrical Engineering, M.I.T., 1948.

mining the channel capacity C can be solved explicitly. However, upper and lower bounds can be set for C in terms of the average noise power N and the noise entropy power N_1 . These bounds are sufficiently close together in most practical cases to furnish a satisfactory solution to the problem.

Theorem 18: The capacity of a channel of band W perturbed by an arbitrary noise is bounded by the inequalities

$$W \log \frac{P + N_1}{N_1} \leq C \leq W \log \frac{P + N}{N_1}$$

where

P = average transmitter power

N = average noise power

N_1 = entropy power of the noise.

Here again the average power of the perturbed signals will be $P + N$. The maximum entropy for this power would occur if the received signal were white noise and would be $W \log 2\pi e(P + N)$. It may not be possible to achieve this; i.e. there may not be any ensemble of transmitted signals which, added to the perturbing noise, produce a white thermal noise at the receiver, but at least this sets an upper bound to $H(y)$. We have, therefore

$$\begin{aligned} C &= \max H(y) - H(n) \\ &\leq W \log 2\pi e(P + N) - W \log 2\pi eN_1. \end{aligned}$$

This is the upper limit given in the theorem. The lower limit can be obtained by considering the rate if we make the transmitted signal a white noise, of power P . In this case the entropy power of the received signal must be at least as great as that of a white noise of power $P + N_1$ since we have shown in a previous theorem that the entropy power of the sum of two ensembles is greater than or equal to the sum of the individual entropy powers. Hence

$$\max H(y) \geq W \log 2\pi e(P + N_1)$$

and

$$\begin{aligned} C &\geq W \log 2\pi e(P + N_1) - W \log 2\pi eN_1 \\ &= W \log \frac{P + N_1}{N_1}. \end{aligned}$$

As P increases, the upper and lower bounds approach each other, so we have as an asymptotic rate

$$W \log \frac{P + N}{N_1}$$

If the noise is itself white, $N = N_1$ and the result reduces to the formula proved previously:

$$C = W \log \left(1 + \frac{P}{N} \right).$$

If the noise is Gaussian but with a spectrum which is not necessarily flat, N_1 is the geometric mean of the noise power over the various frequencies in the band W . Thus

$$N_1 = \exp \frac{1}{W} \int_w \log N(f) df$$

where $N(f)$ is the noise power at frequency f .

Theorem 19: If we set the capacity for a given transmitter power P equal to

$$C = W \log \frac{P + N - \eta}{N_1}$$

then η is monotonic decreasing as P increases and approaches 0 as a limit.

Suppose that for a given power P_1 the channel capacity is

$$W \log \frac{P_1 + N - \eta_1}{N_1}$$

This means that the best signal distribution, say $p(x)$, when added to the noise distribution $q(x)$, gives a received distribution $r(y)$ whose entropy power is $(P_1 + N - \eta_1)$. Let us increase the power to $P_1 + \Delta P$ by adding a white noise of power ΔP to the signal. The entropy of the received signal is now at least

$$H(y) = W \log 2\pi e(P_1 + N - \eta_1 + \Delta P)$$

by application of the theorem on the minimum entropy power of a sum. Hence, since we can attain the H indicated, the entropy of the maximizing distribution must be at least as great and η must be monotonic decreasing. To show that $\eta \rightarrow 0$ as $P \rightarrow \infty$ consider a signal which is a white noise with a large P . Whatever the perturbing noise, the received signal will be approximately a white noise, if P is sufficiently large, in the sense of having an entropy power approaching $P + N$.

25. THE CHANNEL CAPACITY WITH A PEAK POWER LIMITATION

In some applications the transmitter is limited not by the average power output but by the peak instantaneous power. The problem of calculating the channel capacity is then that of maximizing (by variation of the ensemble of transmitted symbols)

$$H(y) - H(n)$$

subject to the constraint that all the functions $f(t)$ in the ensemble be less than or equal to $\sqrt{S}$, say, for all t . A constraint of this type does not work out as well mathematically as the average power limitation. The most we have obtained for this case is a lower bound valid for all $\frac{S}{N}$, an "asymptotic" upper band (valid for large $\frac{S}{N}$) and an asymptotic value of C for $\frac{S}{N}$ small.

Theorem 20: The channel capacity C for a band W perturbed by white thermal noise of power N is bounded by

$$C \geq W \log \frac{2}{\pi e^3} \frac{S}{N},$$

where S is the peak allowed transmitter power. For sufficiently large $\frac{S}{N}$

$$C \leq W \log \frac{\frac{2}{\pi e} S + N}{N} (1 + \epsilon)$$

where ϵ is arbitrarily small. As $\frac{S}{N} \rightarrow 0$ (and provided the band W starts at 0)

$$C \rightarrow W \log \left(1 + \frac{S}{N} \right).$$

We wish to maximize the entropy of the received signal. If $\frac{S}{N}$ is large this will occur very nearly when we maximize the entropy of the transmitted ensemble.

The asymptotic upper bound is obtained by relaxing the conditions on the ensemble. Let us suppose that the power is limited to S not at every instant of time, but only at the sample points. The maximum entropy of the transmitted ensemble under these weakened conditions is certainly greater than or equal to that under the original conditions. This altered problem can be solved easily. The maximum entropy occurs if the different samples are independent and have a distribution function which is constant from $-\sqrt{S}$ to $+\sqrt{S}$. The entropy can be calculated as

$$W \log 4S.$$

The received signal will then have an entropy less than

$$W \log (4S + 2\pi e N)(1 + \epsilon)$$

with $\epsilon \rightarrow 0$ as $\frac{S}{N} \rightarrow \infty$ and the channel capacity is obtained by subtracting the entropy of the white noise, $W \log 2\pi eN$

$$W \log (4S + 2\pi eN)(1 + \epsilon) - W \log (2\pi eN) = W \log \frac{\frac{2}{N} S + N}{N} (1 + \epsilon).$$

This is the desired upper bound to the channel capacity.

To obtain a lower bound consider the same ensemble of functions. Let these functions be passed through an ideal filter with a triangular transfer characteristic. The gain is to be unity at frequency 0 and decline linearly down to gain 0 at frequency W . We first show that the output functions of the filter have a peak power limitation S at all times (not just the sample points). First we note that a pulse $\frac{\sin 2\pi Wt}{2\pi Wt}$ going into the filter produces

$$\frac{1}{2} \frac{\sin^2 \pi Wt}{(\pi Wt)^2}$$

in the output. This function is never negative. The input function (in the general case) can be thought of as the sum of a series of shifted functions

$$a \frac{\sin 2\pi Wt}{2\pi Wt}$$

where a , the amplitude of the sample, is not greater than $\sqrt{S}$. Hence the output is the sum of shifted functions of the non-negative form above with the same coefficients. These functions being non-negative, the greatest positive value for any t is obtained when all the coefficients a have their maximum positive values, i.e. $\sqrt{S}$. In this case the input function was a constant of amplitude $\sqrt{S}$ and since the filter has unit gain for D.C., the output is the same. Hence the output ensemble has a peak power S .

The entropy of the output ensemble can be calculated from that of the input ensemble by using the theorem dealing with such a situation. The output entropy is equal to the input entropy plus the geometrical mean gain of the filter;

$$\int_0^W \log G^2 df = \int_0^W \log \left(\frac{W-f}{W} \right)^2 df = -2W$$

Hence the output entropy is

$$W \log 4S - 2W = W \log \frac{4S}{e^2}$$

and the channel capacity is greater than

$$W \log \frac{2}{\pi e^2} \frac{S}{N}.$$

We now wish to show that, for small $\frac{S}{N}$ (peak signal power over average white noise power), the channel capacity is approximately

$$C = W \log \left(1 + \frac{S}{N} \right).$$

More precisely $C/W \log \left(1 + \frac{S}{N} \right) \rightarrow 1$ as $\frac{S}{N} \rightarrow 0$. Since the average signal power P is less than or equal to the peak S , it follows that for all $\frac{S}{N}$

$$C \leq W \log \left(1 + \frac{P}{N} \right) \leq W \log \left(1 + \frac{S}{N} \right).$$

Therefore, if we can find an ensemble of functions such that they correspond to a rate nearly $W \log \left(1 + \frac{S}{N} \right)$ and are limited to band W and peak S the result will be proved. Consider the ensemble of functions of the following type. A series of t samples have the same value, either $+\sqrt{S}$ or $-\sqrt{S}$, then the next t samples have the same value, etc. The value for a series is chosen at random, probability $\frac{1}{2}$ for $+\sqrt{S}$ and $\frac{1}{2}$ for $-\sqrt{S}$. If this ensemble be passed through a filter with triangular gain characteristic (unit gain at D.C.), the output is peak limited to $\pm S$. Furthermore the average power is nearly S and can be made to approach this by taking t sufficiently large. The entropy of the sum of this and the thermal noise can be found by applying the theorem on the sum of a noise and a small signal. This theorem will apply if

$$\sqrt{t} \frac{S}{N}$$

is sufficiently small. This can be insured by taking $\frac{S}{N}$ small enough (after t is chosen). The entropy power will be $S + N$ to as close an approximation as desired, and hence the rate of transmission as near as we wish to

$$W \log \left(\frac{S + N}{N} \right).$$

PART V: THE RATE FOR A CONTINUOUS SOURCE

26. FIDELITY EVALUATION FUNCTIONS

In the case of a discrete source of information we were able to determine a definite rate of generating information, namely the entropy of the underlying stochastic process. With a continuous source the situation is considerably more involved. In the first place a continuously variable quantity can assume an infinite number of values and requires, therefore, an infinite number of binary digits for exact specification. This means that to transmit the output of a continuous source with *exact recovery* at the receiving point requires, in general, a channel of infinite capacity (in bits per second). Since, ordinarily, channels have a certain amount of noise, and therefore a finite capacity, exact transmission is impossible.

This, however, evades the real issue. Practically, we are not interested in exact transmission when we have a continuous source, but only in transmission to within a certain tolerance. The question is, can we assign a definite rate to a continuous source when we require only a certain fidelity of recovery, measured in a suitable way. Of course, as the fidelity requirements are increased the rate will increase. It will be shown that we can, in very general cases, define such a rate, having the property that it is possible, by properly encoding the information, to transmit it over a channel whose capacity is equal to the rate in question, and satisfy the fidelity requirements. A channel of smaller capacity is insufficient.

It is first necessary to give a general mathematical formulation of the idea of fidelity of transmission. Consider the set of messages of a long duration, say T seconds. The source is described by giving the probability density, in the associated space, that the source will select the message in question $P(x)$. A given communication system is described (from the external point of view) by giving the conditional probability $P_z(y)$ that if message x is produced by the source the recovered message at the receiving point will be y . The system as a whole (including source and transmission system) is described by the probability function $P(x, y)$ of having message x and final output y . If this function is known, the complete characteristics of the system from the point of view of fidelity are known. Any evaluation of fidelity must correspond mathematically to an operation applied to $P(x, y)$. This operation must at least have the properties of a simple ordering of systems; i.e. it must be possible to say of two systems represented by $P_1(x, y)$ and $P_2(x, y)$ that, according to our fidelity criterion, either (1) the first has higher fidelity, (2) the second has higher fidelity, or (3) they have

equal fidelity. This means that a criterion of fidelity can be represented by a numerically valued function:

$$v(P(x, y))$$

whose argument ranges over possible probability functions $P(x, y)$.

We will now show that under very general and reasonable assumptions the function $v(P(x, y))$ can be written in a seemingly much more specialized form, namely as an average of a function $\rho(x, y)$ over the set of possible values of x and y :

$$v(P(x, y)) = \iint P(x, y) \rho(x, y) dx dy$$

To obtain this we need only assume (1) that the source and system are ergodic so that a very long sample will be, with probability nearly 1, typical of the ensemble, and (2) that the evaluation is "reasonable" in the sense that it is possible, by observing a typical input and output x_1 and y_1 , to form a tentative evaluation on the basis of these samples; and if these samples are increased in duration the tentative evaluation will, with probability 1, approach the exact evaluation based on a full knowledge of $P(x, y)$. Let the tentative evaluation be $\rho(x, y)$. Then the function $\rho(x, y)$ approaches (as $T \rightarrow \infty$) a constant for almost all (x, y) which are in the high probability region corresponding to the system:

$$\rho(x, y) \rightarrow v(P(x, y))$$

and we may also write

$$\rho(x, y) \rightarrow \iint P(x, y) \rho(x, y) dx, dy$$

since

$$\iint P(x, y) dx dy = 1$$

This establishes the desired result.

The function $\rho(x, y)$ has the general nature of a "distance" between x and y .⁹ It measures how bad it is (according to our fidelity criterion) to receive y when x is transmitted. The general result given above can be restated as follows: Any reasonable evaluation can be represented as an average of a distance function over the set of messages and recovered messages x and y weighted according to the probability $P(x, y)$ of getting the pair in question, provided the duration T of the messages be taken sufficiently large.

⁹ It is not a "metric" in the strict sense, however, since in general it does not satisfy either $\rho(x, y) = \rho(y, x)$ or $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$.

The following are simple examples of evaluation functions:

1. R.M.S. Criterion.

$$v = \overline{(x(t) - y(t))^2}$$

In this very commonly used criterion of fidelity the distance function $\rho(x, y)$ is (apart from a constant factor) the square of the ordinary euclidean distance between the points x and y in the associated function space.

$$\rho(x, y) = \frac{1}{T} \int_0^T [x(t) - y(t)]^2 dt$$

2. Frequency weighted R.M.S. criterion. More generally one can apply different weights to the different frequency components before using an R.M.S. measure of fidelity. This is equivalent to passing the difference $x(t) - y(t)$ through a shaping filter and then determining the average power in the output. Thus let

$$e(t) = x(t) - y(t)$$

and

$$f(t) = \int_{-\infty}^{\infty} e(\tau) k(t - \tau) d\tau$$

then

$$\rho(x, y) = \frac{1}{T} \int_0^T f(t)^2 dt.$$

3. Absolute error criterion.

$$\rho(x, y) = \frac{1}{T} \int_0^T |x(t) - y(t)| dt$$

4. The structure of the ear and brain determine implicitly an evaluation, or rather a number of evaluations, appropriate in the case of speech or music transmission. There is, for example, an "intelligibility" criterion in which $\rho(x, y)$ is equal to the relative frequency of incorrectly interpreted words when message $x(t)$ is received as $y(t)$. Although we cannot give an explicit representation of $\rho(x, y)$ in these cases it could, in principle, be determined by sufficient experimentation. Some of its properties follow from well-known experimental results in hearing, e.g., the ear is relatively insensitive to phase and the sensitivity to amplitude and frequency is roughly logarithmic.
5. The discrete case can be considered as a specialization in which we have

tacitly assumed an evaluation based on the frequency of errors. The function $\rho(x, y)$ is then defined as the number of symbols in the sequence y differing from the corresponding symbols in x divided by the total number of symbols in x .

27. THE RATE FOR A SOURCE RELATIVE TO A FIDELITY EVALUATION

We are now in a position to define a rate of generating information for a continuous source. We are given $P(x)$ for the source and an evaluation v determined by a distance function $\rho(x, y)$ which will be assumed continuous in both x and y . With a particular system $P(x, y)$ the quality is measured by

$$v = \iint \rho(x, y) P(x, y) dx dy$$

Furthermore the rate of flow of binary digits corresponding to $P(x, y)$ is

$$R = \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy.$$

We define the rate R_1 of generating information for a given quality v_1 of reproduction to be the minimum of R when we keep v fixed at v_1 and vary $P_x(y)$. That is:

$$R_1 = \text{Min}_{P_x(y)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

subject to the constraint:

$$v_1 = \iint P(x, y)\rho(x, y) dx dy.$$

This means that we consider, in effect, all the communication systems that might be used and that transmit with the required fidelity. The rate of transmission in bits per second is calculated for each one and we choose that having the least rate. This latter rate is the rate we assign the source for the fidelity in question.

The justification of this definition lies in the following result:

Theorem 21: If a source has a rate R_1 for a valuation v_1 it is possible to encode the output of the source and transmit it over a channel of capacity C with fidelity as near v_1 as desired provided $R_1 \leq C$. This is not possible if $R_1 > C$.

The last statement in the theorem follows immediately from the definition of R_1 and previous results. If it were not true we could transmit more than C bits per second over a channel of capacity C . The first part of the theorem is proved by a method analogous to that used for Theorem 11. We may, in the first place, divide the (x, y) space into a large number of small cells and

represent the situation as a discrete case. This will not change the evaluation function by more than an arbitrarily small amount (when the cells are very small) because of the continuity assumed for $\rho(x, y)$. Suppose that $P_1(x, y)$ is the particular system which minimizes the rate and gives R_1 . We choose from the high probability y 's a set at random containing

$$2^{(R_1 + \epsilon)T}$$

members where $\epsilon \rightarrow 0$ as $T \rightarrow \infty$. With large T each chosen point will be connected by a high probability line (as in Fig. 10) to a set of x 's. A calculation similar to that used in proving Theorem 11 shows that with large T almost all x 's are covered by the fans from the chosen y points for almost all choices of the y 's. The communication system to be used operates as follows: The selected points are assigned binary numbers. When a message x is originated it will (with probability approaching 1 as $T \rightarrow \infty$) lie within one at least of the fans. The corresponding binary number is transmitted (or one of them chosen arbitrarily if there are several) over the channel by suitable coding means to give a small probability of error. Since $R_1 \leq C$ this is possible. At the receiving point the corresponding y is reconstructed and used as the recovered message.

The evaluation v_1' for this system can be made arbitrarily close to v_1 by taking T sufficiently large. This is due to the fact that for each long sample of message $x(t)$ and recovered message $y(t)$ the evaluation approaches v_1 (with probability 1).

It is interesting to note that, in this system, the noise in the recovered message is actually produced by a kind of general quantizing at the transmitter and is not produced by the noise in the channel. It is more or less analogous to the quantizing noise in P.C.M.

28. THE CALCULATION OF RATES

The definition of the rate is similar in many respects to the definition of channel capacity. In the former

$$R = \text{Max}_{P_x(y)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

with $P(x)$ and $v_1 = \iint P(x, y)\rho(x, y) dx dy$ fixed. In the latter

$$C = \text{Min}_{P(x)} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy$$

with $P_x(y)$ fixed and possibly one or more other constraints (e.g., an average power limitation) of the form $K = \iint P(x, y) \lambda(x, y) dx dy$.

A partial solution of the general maximizing problem for determining the rate of a source can be given. Using Lagrange's method we consider

$$\iint \left[P(x, y) \log \frac{P(x, y)}{P(x)P(y)} + \mu P(x, y)\rho(x, y) + \nu(x)P(x, y) \right] dx dy$$

The variational equation (when we take the first variation on $P(x, y)$) leads to

$$P_y(x) = B(x) e^{-\lambda \rho(x, y)}$$

where λ is determined to give the required fidelity and $B(x)$ is chosen to satisfy

$$\int B(x) e^{-\lambda \rho(x, y)} dx = 1$$

This shows that, with best encoding, the conditional probability of a certain cause for various received y , $P_y(x)$ will decline exponentially with the distance function $\rho(x, y)$ between the x and y is question.

In the special case where the distance function $\rho(x, y)$ depends only on the (vector) difference between x and y ,

$$\rho(x, y) = \rho(x - y)^*$$

we have

$$\int B(x) e^{-\lambda \rho(x-y)} dx = 1.$$

Hence $B(x)$ is constant, say α , and

$$P_y(x) = \alpha e^{-\lambda \rho(x-y)}$$

Unfortunately these formal solutions are difficult to evaluate in particular cases and seem to be of little value. In fact, the actual calculation of rates has been carried out in only a few very simple cases.

If the distance function $\rho(x, y)$ is the mean square discrepancy between x and y and the message ensemble is white noise, the rate can be determined. In that case we have

$$R = \text{Min} [H(x) - H_y(x)] = H(x) - \text{Max} H_y(x)$$

with $N = \overline{(x - y)^2}$. But the $\text{Max} H_y(x)$ occurs when $y - x$ is a white noise, and is equal to $W_1 \log 2\pi e N$ where W_1 is the bandwidth of the message ensemble. Therefore

$$\begin{aligned} R &= W_1 \log 2\pi e Q - W_1 \log 2\pi e N \\ &= W_1 \log \frac{Q}{N} \end{aligned}$$

where Q is the average message power. This proves the following:

Theorem 22: The rate for a white noise source of power Q and band W_1 relative to an R.M.S. measure of fidelity is

$$R = W_1 \log \frac{Q}{N}$$

where N is the allowed mean square error between original and recovered messages.

More generally with any message source we can obtain inequalities bounding the rate relative to a mean square error criterion.

Theorem 23: The rate for any source of band W_1 is bounded by

$$W_1 \log \frac{Q_1}{N} \leq R \leq W_1 \log \frac{Q}{N}$$

where Q is the average power of the source, Q_1 its entropy power and N the allowed mean square error.

The lower bound follows from the fact that the $\max H_y(x)$ for a given $(x - y)^2 = N$ occurs in the white noise case. The upper bound results if we place the points (used in the proof of Theorem 21) not in the best way but at random in a sphere of radius $\sqrt{Q - N}$.

ACKNOWLEDGMENTS

The writer is indebted to his colleagues at the Laboratories, particularly to Dr. H. W. Bode, Dr. J. R. Pierce, Dr. B. McMillan, and Dr. B. M. Oliver for many helpful suggestions and criticisms during the course of this work. Credit should also be given to Professor N. Wiener, whose elegant solution of the problems of filtering and prediction of stationary ensembles has considerably influenced the writer's thinking in this field.

APPENDIX 5

Let S_1 be any measurable subset of the g ensemble, and S_2 the subset of the f ensemble which gives S_1 under the operation T . Then

$$S_1 = TS_2.$$

Let H^λ be the operator which shifts all functions in a set by the time λ . Then

$$H^\lambda S_1 = H^\lambda TS_2 = TH^\lambda S_2$$

since T is invariant and therefore commutes with H^λ . Hence if $m[S]$ is the probability measure of the set S

$$\begin{aligned} m[H^\lambda S_1] &= m[TH^\lambda S_2] = m[H^\lambda S_2] \\ &= m[S_2] = m[S_1] \end{aligned}$$

where the second equality is by definition of measure in the g space the third since the f ensemble is stationary, and the last by definition of g measure again.

To prove that the ergodic property is preserved under invariant operations, let S_1 be a subset of the g ensemble which is invariant under H^λ , and let S_2 be the set of all functions f which transform into S_1 . Then

$$H^\lambda S_1 = H^\lambda T S_2 = T H^\lambda S_2 = S_1$$

so that $H^\lambda S_1$ is included in S_1 for all λ . Now, since

$$m[H^\lambda S_2] = m[S_1]$$

this implies

$$H^\lambda S_2 = S_2$$

for all λ with $m[S_2] \neq 0, 1$. This contradiction shows that S_1 does not exist.

APPENDIX 6

The upper bound, $\bar{N}_3 \leq N_1 + N_2$, is due to the fact that the maximum possible entropy for a power $N_1 + N_2$ occurs when we have a white noise of this power. In this case the entropy power is $N_1 + N_2$.

To obtain the lower bound, suppose we have two distributions in n dimensions $p(x_i)$ and $q(x_i)$ with entropy powers $\bar{N}_1$ and $\bar{N}_2$. What form should p and q have to minimize the entropy power $\bar{N}_3$ of their convolution $r(x_i)$:

$$r(x_i) = \int p(y_i) q(x_i - y_i) dy_i.$$

The entropy H_3 of r is given by

$$H_3 = - \int r(x_i) \log r(x_i) dx_i.$$

We wish to minimize this subject to the constraints

$$H_1 = - \int p(x_i) \log p(x_i) dx_i$$

$$H_2 = - \int q(x_i) \log q(x_i) dx_i.$$

We consider then

$$U = - \int [r(x) \log r(x) + \lambda p(x) \log p(x) + \mu q(x) \log q(x)] dx$$

$$\delta U = - \int [[1 + \log r(x)] \delta r(x) + \lambda [1 + \log p(x)] \delta p(x)$$

$$+ \mu [1 + \log q(x)] \delta q(x)] dx.$$

If $p(x)$ is varied at a particular argument $x_i = s_i$, the variation in $r(x)$ is

$$\delta r(x) = q(x_i - s_i)$$

and

$$\delta U = - \int q(x_i - s_i) \log r(x_i) dx_i - \lambda \log p(s_i) = 0$$

and similarly when q is varied. Hence the conditions for a minimum are

$$\int q(x_i - s_i) \log r(x_i) = -\lambda \log p(s_i)$$

$$\int p(x_i - s_i) \log r(x_i) = -\mu \log q(s_i).$$

If we multiply the first by $p(s_i)$ and the second by $q(s_i)$ and integrate with respect to s we obtain

$$H_3 = -\lambda H_1$$

$$H_3 = -\mu H_2$$

or solving for λ and μ and replacing in the equations

$$H_1 \int q(x_i - s_i) \log r(x_i) dx_i = -H_3 \log p(s_i)$$

$$H_2 \int p(x_i - s_i) \log r(x_i) dx_i = -H_3 \log p(s_i).$$

Now suppose $p(x_i)$ and $q(x_i)$ are normal

$$p(x_i) = \frac{|A_{ij}|^{n/2}}{(2\pi)^{n/2}} \exp - \frac{1}{2} \Sigma A_{ij} x_i x_j$$

$$q(x_i) = \frac{|B_{ij}|^{n/2}}{(2\pi)^{n/2}} \exp - \frac{1}{2} \Sigma B_{ij} x_i x_j.$$

Then $r(x_i)$ will also be normal with quadratic form C_{ij} . If the inverses of these forms are a_{ij} , b_{ij} , c_{ij} then

$$c_{ij} = a_{ij} + b_{ij}.$$

We wish to show that these functions satisfy the minimizing conditions if and only if $a_{ij} = K b_{ij}$ and thus give the minimum H_3 under the constraints. First we have

$$\log r(x_i) = \frac{n}{2} \log \frac{1}{2\pi} |C_{ij}| - \frac{1}{2} \Sigma C_{ij} x_i x_j$$

$$\int q(x_i - s_i) \log r(x_i) = \frac{n}{2} \log \frac{1}{2\pi} |C_{ij}| - \frac{1}{2} \Sigma C_{ij} s_i s_j - \frac{1}{2} \Sigma C_{ij} b_{ij}.$$

This should equal

$$\frac{H_3}{H_1} \left[\frac{n}{2} \log \frac{1}{2\pi} |A_{ij}| - \frac{1}{2} \Sigma A_{ij} s_i s_j \right]$$

which requires $A_{ij} = \frac{H_1}{H_3} C_{ij}$.

In this case $A_{ij} = \frac{H_1}{H_2} B_{ij}$ and both equations reduce to identities.

APPENDIX 7

The following will indicate a more general and more rigorous approach to the central definitions of communication theory. Consider a probability measure space whose elements are ordered pairs (x, y) . The variables x, y are to be identified as the possible transmitted and received signals of some long duration T . Let us call the set of all points whose x belongs to a subset S_1 of x points the strip over S_1 , and similarly the set whose y belongs to S_2 the strip over S_2 . We divide x and y into a collection of non-overlapping measurable subsets X_i and Y_i approximate to the rate of transmission R by

$$R_1 = \frac{1}{T} \sum_i P(X_i, Y_i) \log \frac{P(X_i, Y_i)}{P(X_i)P(Y_i)}$$

where

$P(X_i)$ is the probability measure of the strip over X_i

$P(Y_i)$ is the probability measure of the strip over Y_i

$P(X_i, Y_i)$ is the probability measure of the intersection of the strips.

A further subdivision can never decrease R_1 . For let X_1 be divided into $X_1 = X'_1 + X''_1$ and let

$$P(Y_1) = a \qquad P(X_1) = b + c$$

$$P(X'_1) = b \qquad P(X'_1, Y_1) = d$$

$$P(X''_1) = c \qquad P(X''_1, Y_1) = e$$

$$P(X_1, Y_1) = d + e$$

Then in the sum we have replaced (for the X_1, Y_1 intersection)

$$(d + e) \log \frac{d + e}{a(b + c)} \quad \text{by} \quad d \log \frac{d}{ab} + e \log \frac{e}{ac}.$$

It is easily shown that with the limitation we have on b, c, d, e ,

$$\left[\frac{d + e}{b + c} \right]^{d+e} \leq \frac{d^d e^e}{b^d c^e}$$

and consequently the sum is increased. Thus the various possible subdivisions form a directed set, with R monotonic increasing with refinement of the subdivision. We may define R unambiguously as the least upper bound for the R_1 and write it

$$R = \frac{1}{T} \iint P(x, y) \log \frac{P(x, y)}{P(x)P(y)} dx dy.$$

This integral, understood in the above sense, includes both the continuous and discrete cases and of course many others which cannot be represented in either form. It is trivial in this formulation that if x and u are in one-to-one correspondence, the rate from u to y is equal to that from x to y . If v is any function of y (not necessarily with an inverse) then the rate from x to y is greater than or equal to that from x to v since, in the calculation of the approximations, the subdivisions of y are essentially a finer subdivision of those for v . More generally if y and v are related not functionally but statistically, i.e., we have a probability measure space (y, v) , then $R(x, v) \leq R(x, y)$. This means that any operation applied to the received signal, even though it involves statistical elements, does not increase R .

Another notion which should be defined precisely in an abstract formulation of the theory is that of "dimension rate," that is the average number of dimensions required per second to specify a member of an ensemble. In the band limited case $2W$ numbers per second are sufficient. A general definition can be framed as follows. Let $f_\alpha(t)$ be an ensemble of functions and let $\rho_T[f_\alpha(t), f_\beta(t)]$ be a metric measuring the "distance" from f_α to f_β over the time T (for example the R.M.S. discrepancy over this interval.) Let $N(\epsilon, \delta, T)$ be the least number of elements f which can be chosen such that all elements of the ensemble apart from a set of measure δ are within the distance ϵ of at least one of those chosen. Thus we are covering the space to within ϵ apart from a set of small measure δ . We define the dimension rate λ for the ensemble by the triple limit

$$\lambda = \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{\log N(\epsilon, \delta, T)}{T \log \epsilon}.$$

This is a generalization of the measure type definitions of dimension in topology, and agrees with the intuitive dimension rate for simple ensembles where the desired result is obvious.

Transients in Mechanical Systems

By J. T. MULLER

INTRODUCTION

A study of the response of an electrical network or system to the input of transients in the form of short-duration pulses is an accepted method of analysis of the network. By comparing the input and the output, conclusions may be drawn as to the respective merit of the various components.

Until recently similar procedures were only of academic interest with mechanical systems. However, the tests for mechanical ruggedness, which are required of electronic gear in order to pass specifications for the armed forces, are an example of the application of transients to a mechanical system. These tests are known as *High Impact Shock Tests*.

A basic part of an electrical system is a damped resonant network consisting of an inductance, a capacitance and a resistance. A mass, a spring and a friction device is the equivalent mechanical network called a simple mechanical system and a combination of such networks is a general mechanical system. It is, of course, advantageous to keep the mechanical system as simple as possible without detracting from the general usefulness of the results obtained.

The problems here considered are pertinent to a system which is essentially made up of a supporting structure or table and a resilient mounting array bearing the equipment (e.g. electronic gear) which is vulnerable to shock. (See Fig. 1.)

A shock is the physical manifestation of the transfer of mechanical energy from one body to another during an extremely short interval of time. The order of magnitude of the time interval is milliseconds and quite frequently fractions of a millisecond.

The system is excited by administering large spurts of mechanical energy to the supporting table. The manner in which this energy is supplied to the base and the way it is dissipated through the system are the subjects of this paper.

The energy transfer to the supporting table is accomplished by the use of huge hammers which strike the anvil with controllable speeds. The action is assumed to be similar to that of an explosion, particularly to an underwater explosion at close range or a *near-miss*. As to the real comparison between the two, the reader is referred to the various manuscripts published by the Bureau of Ships. This particular phase of the subject is

considered outside the scope of this paper, except for the following brief statement:

Both actions fit the definition of shock stated above and the difference between the two is one of size and not of kind.

Shocks are transients and are conveniently treated by a branch of mathematics which is adapted to the solution of problems of this kind; viz, the *Laplace Transforms*, and the reader is referred to Gardner and Barnes, "Transients in Linear Systems." The nomenclature used here is identical to that of those authors.

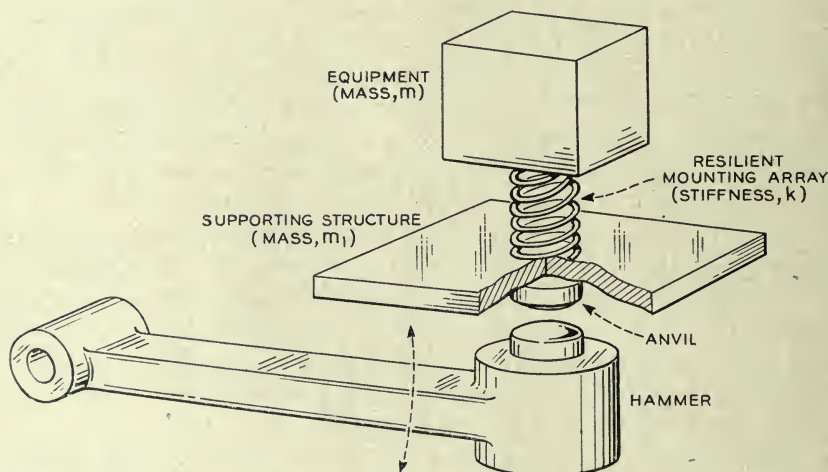


Fig. 1—Schematic layout of shock machine.

The manuscript consists of two parts: In the first, the energy transfer to the base is considered. We are dealing here with rigid bodies; consequently with very small transient displacements and very large forces. These are usually referred to as impact forces or impulses and four such functions of force and time are discussed. Displacements with associated velocities result from the action of impulses on the base.

The second part deals with the effect of these displacements on the shock-mounted equipment. Although the mathematical procedure is identical to the first part, here we deal with a function of displacement and time. There is no specific name for such a relationship but a suggestive term is "whip." However, the pulse functions represented are the same as those of the force and time function.

It is assumed that the displacement-time pulse is independent of the subsequent motion of the mass.

In considering any kind of shock problem we have the following fundamental considerations:

First, we shall want to know the magnitude of the shock present in the base or supporting structure; this will be called the "excitation."

Second, the behavior of the resilient medium interposed between the shock-producing base and the equipment. It is sometimes expressed as the coupling. We shall use the term *transmission*.

Third, the resulting disturbance of the equipment caused by the transmitted shock, which we will call the *response*.

The three functions do not exist independently, but are mathematically related. For a clearer understanding of shock phenomena it is perhaps helpful to fix in one's mind the idea that the response of a system is completely dependent upon the transmission function.

To use an electrical analogue, the voltage $e_1(t)$ impressed upon a system produces an output voltage $e_2(t)$ which is completely defined by the transmission function. For instance, if this transmission function represents a filter of some kind with given boundaries, then it is to be expected that the response of $e_2(t)$ is completely changed outside these limits and could even be zero. The same train of thought will hold for mechanical systems. Here the transmission function is mostly represented by the stiffness or the compliance. For a completely rigid medium the stiffness would be infinite and the input and output would be alike; in other words, a force applied to the base would appear at the equipment. This is a theoretical case because no material is perfectly rigid. Though some materials are more rigid than others they will all give if the force applied is big enough. Now the forces associated with a shock are almost always of considerable magnitude so that the stiffness of a material becomes significant.

As the stiffness diminishes the response changes and may appear to be quite different from the input. As far as the transmissibility of forces is concerned, the reader is reminded that a force is always accompanied by a reaction. The forces which put the base into motion cannot be transmitted by a soft material like rubber, unless it is compressed to extremely high values, and thus produce an equally large reactive force.

PART I

ANALYSIS OF THE EXCITATION OF THE BASE

By recording the motions of the base, we obtain time-displacement curves as shown in Fig. 2. The method of recording has been done by means of high-speed motion pictures (at the Whippany testing laboratory using a Fastex) and by using strain gages (at the Annapolis Engineering Experiment Station).

The test equipments are fundamentally mechanical impact producing machines. For technical details and description of the machines the reader

is referred to the various test specifications by the Bureau of Ships (as, for example, *Spec. 40T9*).

The characteristic of an impact is the transfer of mechanical energy from one mass to another in a relatively short time. The corresponding force as function of time is called an *impulse*, henceforth indicated as $F(t)$. A study of the pulse functions has suggested some probable theoretical shapes of $F(t)$ which could cover a wide variety of conditions. These pulse functions will be used for force-time functions as well as displacement-time-functions and it will be shown that the results are surprisingly similar.

We will let these pulses operate on the base with mass m_1 and calculate and plot the resulting time displacement curves. Since an impulse is associated with energy transfer, it must be a function of $\frac{m_1 v^2}{2}$. From the point of view

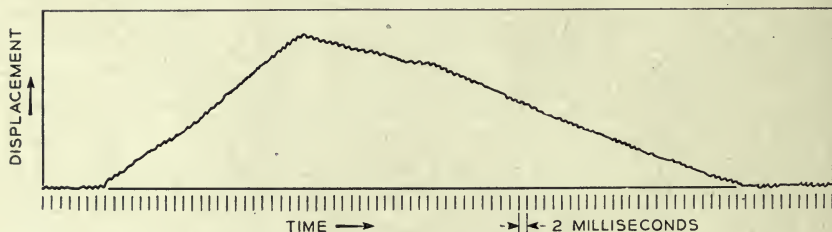


Fig. 2—Time displacement record of medium high impact machine.

of shock action, the final velocity v is extremely important, for it is this velocity which will determine the displacement and acceleration of the shock-mounted equipment.

To distinguish the various applications of the pulse functions, the following notations are adopted:

$f(t)$ represents any functions of t , without reference to its dimensional magnitude. The transform of $f(t)$ is indicated by $F(s)$.

$x(t)$ represents a function of t when it is a displacement of the mass m only. The transform is indicated by $X(s)$.

$x_1(t)$ represents a function of t when it is a displacement of the base (with mass m_1) only. The transform is indicated by $X_1(s)$.

$F(t)$ represents a function of t when it is a force applied to the base. The transform is indicated by $F_0(s)$.

Since $x_1(t)$ and $F(t)$ are input functions, they may be represented by the same type pulse, in which case the transforms are alike, i.e., $F(s) = X_1(s) = F_0(s)$.

Figure 3A, a rectangular pulse, is the simplest form.

Figure 3B is a triangular pulse, $f(t)$, reaching a peak and returning to zero in a linear manner.

Figure 3C consists of one-half cycle of a sine wave.

Figure 3D is a cosine pulse of one-cycle duration and is shifted along the Y axis an amount equal to the amplitude.

These are the pulses to be used in the problems under consideration.

If they represent a force as it varies with time then it is said that $F(t)$ represents a particular pulse. The Laplace transform of $F(t)$ is given as $F_0(s)$, $F_0(s)$ being some function in the complex domain. It is outside the scope of this paper to prove or show the mathematical technique in obtaining the transforms which produce $F_0(s)$. We will present them here for future reference.

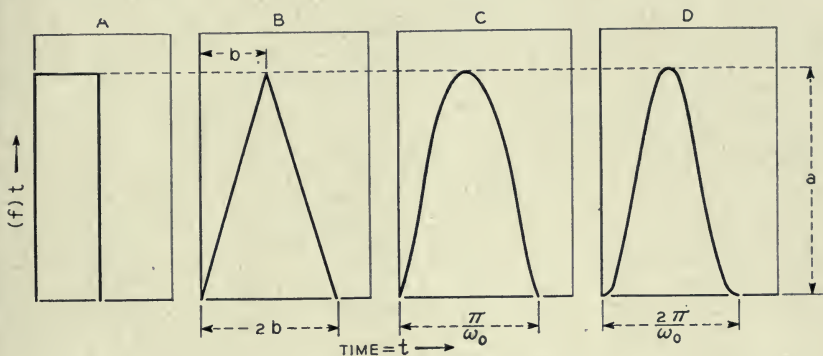


Fig. 3—Four pulses.

The Laplace transform for a very short pulse is

$$F(s) = A_r \quad (1.01)$$

and is a pulse which has a finite area but the time interval of which is approaching zero.

For a square pulse with finite time interval and magnitude a (hereafter referred to as pulse amplitude) it is

$$F(s) = a \frac{1 - e^{-bs}}{s} \quad (1.02)$$

For a triangular pulse

$$F(s) = \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2 \quad (1.03)$$

For a sine pulse

$$F(s) = \frac{a\omega_0}{s^2 + \omega_0^2} (1 + e^{-\pi s/\omega_0}) \quad (1.04)$$

For a shifted cosine pulse

$$F(s) = \frac{a\omega^2}{2s(s^2 + \omega_0^2)} (1 - e^{-2\pi s/\omega_0}) \quad (1.05)$$

Suppose we let an impulse, and to take a specific example, a triangular impulse, operate on a mass m_1 . We have

$$F = m_1 a_0 \quad (1.06)$$

in which

$$F = F(t) = \text{force in lbs.}$$

$$m_1 = \text{mass in slugs}$$

$$a_0 = \ddot{x}_1 = \text{acceleration in ft/sec}^2$$

Let $\mathcal{L}[x_1(t)] = X_1(s) = X_1$
then

$$\mathcal{L}[m_1 \ddot{x}_1] = m_1 s^2 X_1(s)$$

the $\mathcal{L}$ transform of a triangular pulse $F(t)$ is

$$\mathcal{L}[F(t)] = \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2 = F_0(s) \quad (1.07)$$

Substituting

$$X_1 = \frac{a}{m_1 b} \frac{1}{s^2} \left(\frac{1 - e^{-bs}}{s} \right)^2$$

The inverse transform is

$$\mathcal{L}^{-1}[X_1(s)] = x_1(t) = \mathcal{L}^{-1} \left[\frac{a}{m_1 b} \frac{1}{s^2} \left(\frac{1 - e^{-bs}}{s} \right)^2 \right] \quad (1.08)$$

$$x_1(t) = \frac{a}{m_1 b} \mathcal{L}^{-1} \left[\frac{1}{s^2} \left(\frac{1 - e^{-bs}}{s} \right)^2 \right] \quad (1.09)$$

The solution of 1.09

$$x_1(t) = \frac{a}{m_1 b} \left[\frac{t^3}{6} - 2 \frac{(t-b)^3}{6} u(t-b) + \frac{(t-2b)^3}{6} u(t-2b) \right] \quad (1.10)$$

After the impulse is over, i.e., for values of $t > 2b$, 1.10 becomes

$$x_1(t) = \frac{a}{m_1} b(t-b) \quad (1.11)$$

and the final velocity is

$$\frac{a}{m_1} b, \quad (1.12)$$

which represents the area of the impulse divided by the mass.

Similarly we find for the very short square impulse

$$x_1 = \frac{A_r}{m_1} t \quad (1.13)$$

For the square impulse of finite time interval b

$$x_1 = \frac{ab}{m_1} \left(t - \frac{b}{2} \right) \quad (1.14)$$

For the sine impulse

$$x_1 = \frac{2a}{m_1 \omega_0} \left(t - \frac{\pi}{2\omega_0} \right) \quad (1.15)$$

For the shifted cosine impulse

$$x_1 = \frac{a\pi}{m_1 \omega_0} \left(t - \frac{\pi}{\omega_0} \right) \quad (1.16)$$

The velocity is the term preceding the term in parenthesis.

In the five examples mentioned, we find that this *velocity is proportional to the area of the impulse curve and inversely proportional to the mass*. All expressions contain the factor $\frac{a}{m_1}$ and, since a is the maximum force present, this expression represents the maximum acceleration and it is this value which is so frequently mentioned when discussing the actions on the shock table.

For instance, from records we have determined approximate values for the time interval during which the energy transfer from hammer to the table takes place. The high-speed motion pictures are taken at the rate of 4,000 to 5,000 frames per second, which means an average elapsed time of .22 milliseconds or 220 microseconds. The energy transfer occurs within this time interval, because the rate of increase of the displacement from frame to frame is constant. The exposure time of one frame is $\frac{1}{12000}$ second or 83 microseconds. If the anvil moved within this time there would be evidence of blurring. Since we have been unable to detect any blurring, we may state that transfer is less than 220 μ s yet more than 80 μ s.

Let us assume it to be 100 μ s. That means a pulse width of $2b = 100 \mu$ s. (See Fig. 3.)

If

$$\frac{a}{m_1} = a_0 = \text{acceleration,}$$

then

$$v = \frac{ab}{m_1} = a_0 b$$

For a 2000-ft. pound shock the table speed v is approximately 7 ft./sec. Substituting, we find for a_0 or the acceleration

$$7 = a_0 \times .00005$$

$$\text{or } a_0 = 140,000 \text{ ft./sec.}^2$$

$$a_0 = 4400 \text{ "g"s}$$

This is about the order of magnitude which the accelerometers have recorded.

The important conclusion we draw from this is that the acceleration and its time interval combine to produce a velocity of the base which is a complete criterion of the severity of the shock administered.

In the example just cited the weight of the table is approximately 800 lbs., and the force $4400 \times 800 = 3,520,000$ lbs. The result is, then, that a triangular impulse of 3,520,000 lbs. magnitude and a duration of $100\mu\text{s}$ operating on a table of 800 lbs., imparts to that table a velocity of 7 ft./sec.

PART II

ANALYSIS OF THE RESPONSE

In Part I the origin of the motion of the base has been treated. This motion of the base can now be represented by a pulse or a displacement as a function of time. To distinguish the displacement-time function from the force-time function, we have already suggested the name *Whip*. Obviously some of the pulse shapes which were used to represent impulses are not suitable as whips. For instance, the square pulse as whip could not exist, since this would suppose an infinite velocity.

The triangular whip is observed in the medium-high-impact shock machine. The sine whip may be taken to represent approximately the output of the light-high-impact machine.

The shifted cosine whip is sometimes used in the motion of cams of automatic equipment.

The problem of shock response is now reduced to the behavior of a mass and spring system when the base motion is represented by a whip.

Triangular whip (Fig. 3b). The Laplace transform of this pulse is

$$F(s) = \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2.$$

The differential equation for a simple harmonic system is

$$m\ddot{x} + kx = 0 \quad (1)$$

or

$$\frac{\ddot{x}}{\omega^2} + x = 0. \quad (2)$$

If we let the whip operate on this system, then

$$\frac{\ddot{x}}{\omega^2} + x = f(t) = x_1(t) \quad (3)$$

in which $x_1(t)$ represents the displacement of the whip as a function of time.

$$\text{Let } \mathfrak{L}[x(t)] = X(s)$$

and

$$\mathfrak{L}[x_1(t)] = \mathfrak{L}[f(t)] = F(s) = X_1(s)$$

then

$$\mathfrak{L}[\ddot{x}] = s^2 X(s) - sf(0) - f'(0)$$

By definition the initial conditions are zero, so that

$$\mathfrak{L}[\ddot{x}] = s^2 X(s) \quad (4)$$

The Laplace transform of equation (3) is then

$$\frac{s^2}{\omega^2} X(s) + X(s) = \mathfrak{L}[f(t)] = X_1(s)$$

or

$$\left(\frac{s^2 + \omega^2}{\omega^2} \right) X(s) = X_1(s). \quad (5)$$

Now

$$X_1(s) = \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2.$$

Substituting and rearranging,

$$X(s) = \frac{\omega^2}{s^2 + \omega^2} \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2. \quad (6)$$

This is the transform equation. To find x we use the inverse Laplace transform and the solution of (7) is

$$x = \frac{a\omega^2}{b} \left[\left(\frac{t}{\omega^2} - \frac{\sin \omega t}{\omega^3} \right) - 2 \left(\frac{(t-b)}{\omega^2} - \frac{\sin \omega(t-b)}{\omega^3} \right) u(t-b) + \left(\frac{t-2b}{\omega^2} - \frac{\sin \omega(t-2b)}{\omega^3} \right) u(t-2b) \right]. \quad (7)$$

The expression $u(t-b)$ simply means that the term to which it is attached is zero for all values of $t < b$.

Let us now consider what this solution consists of.

There are apparently three terms which take effect at successive intervals.

The initial whip can be considered to consist of three different displacements starting at successive times 0, b and $2b$. With the displacement of the base there is a corresponding displacement of the mass m . After the time b the second term or displacement takes hold and an associated displacement of mass m except that the initial conditions are the end conditions of the first displacement. After the time $2b$ the third displacement enters and the final result is the displacement-time pulse or whip. To make the problem somewhat simpler we introduce the following modifications:

1°. Because the motion is a simple harmonic of known frequency after the whip has passed we will only consider the maximum amplitude.

2°. Only the displacement-time function of the mass m during the pulse interval will be examined.

3°. The dimensional magnitudes of the motion of mass m will be expressed as ratios of those of the pulse.

If a is the maximum amplitude of the whip, and $T_0 = 2b$ its time interval (usually expressed in milliseconds), then we define

$$\frac{x}{a} = \delta \quad \text{Amplitude ratio of pulse displacement and response of mass } m \text{ during pulse interval only.}$$

$$\frac{T_0}{T} = \frac{2b}{T} = \frac{2b}{2\pi/\omega} \frac{\omega b}{\pi} = \varphi \quad \text{Natural frequency of mass } m \text{ expressed as a ratio of the pulse length.}$$

$$\tau = \frac{t}{2b} \quad \text{Elapsed time expressed as a ratio of the pulse length.}$$

$$\Delta = \frac{x_{\max}}{a} \quad \text{Ratio of maximum amplitude to pulse displacement after pulse interval.}$$

Substituting these values in equation (7) and rearranging we obtain

$$\delta = 2\tau - \frac{\sin 2\varphi\pi\tau}{\pi\varphi} - 2 \left((2\tau - 1) - \frac{\sin \pi\varphi(2\tau - 1)}{\pi\varphi} \right) u(2\tau - 1) + \left(2(\tau - 1) - \frac{\sin 2\pi\varphi(\tau - 1)}{\pi\varphi} \right) u(\tau - 1) \quad (8)$$

This looks somewhat complicated, but we can simplify by omitting the last term, because we are only considering values of δ during the pulse interval.

$$\therefore \delta = 2\tau - \frac{\sin 2\varphi\pi\tau}{\pi\varphi} - 2 \left((2\tau - 1) - \frac{\sin \varphi(2\tau - 1)}{\pi\varphi} \right) u(2\tau - 1) \quad (9)$$

A plot of this equation for various values of φ is shown in Fig. 4. It is seen that δ becomes a maximum when φ is approx. .9 and τ is then .75. The displacement is approximately 1.5 times the peak displacement of the whip.

After the whip has passed, or when $\tau > 1$, the transient has disappeared and a steady-state condition exists. Since the system under consideration is a simple harmonic system, the steady state is a harmonic motion of frequency ω , with an amplitude to be obtained from equation (8). Indicating the dimensionless values of the amplitude by δ_a when $\tau > 1$, equation 8 may be written

$$\begin{aligned} \delta_a = 2\tau - \frac{\sin 2\varphi\pi\tau}{\pi\varphi} - 2 \left((2\tau - 1) - \frac{\sin \varphi(2\tau - 1)}{\pi\varphi} \right) \\ + \left(2(\tau - 1) - \frac{\sin 2\pi\varphi(\tau - 1)}{\pi\varphi} \right) \quad \tau > 1 \end{aligned} \quad (10)$$

After developing (10) and rearranging we obtain

$$\delta_a = \frac{2(1 - \cos \pi\varphi)}{\pi\varphi} \sin \pi\varphi(2\tau - 1). \quad (11)$$

The maximum amplitude is

$$\Delta = \frac{2(1 - \cos \pi\varphi)}{\pi\varphi} \quad (12)$$

A plot of equation (12) is shown in Fig. 5. Before considering the action of this whip in terms of what it does to the system, we shall take a brief look at the analysis of the two other whips; viz., the sine whip and shifted cosine whip (see Fig. 3).

Sine Whip

We have again equation (3).

$$\frac{\ddot{x}}{\omega^2} + x = f(t) = x_1(t)$$

and equation (5)

$$\left(\frac{s^2 + \omega^2}{\omega^2} \right) X(s) = F(s)$$

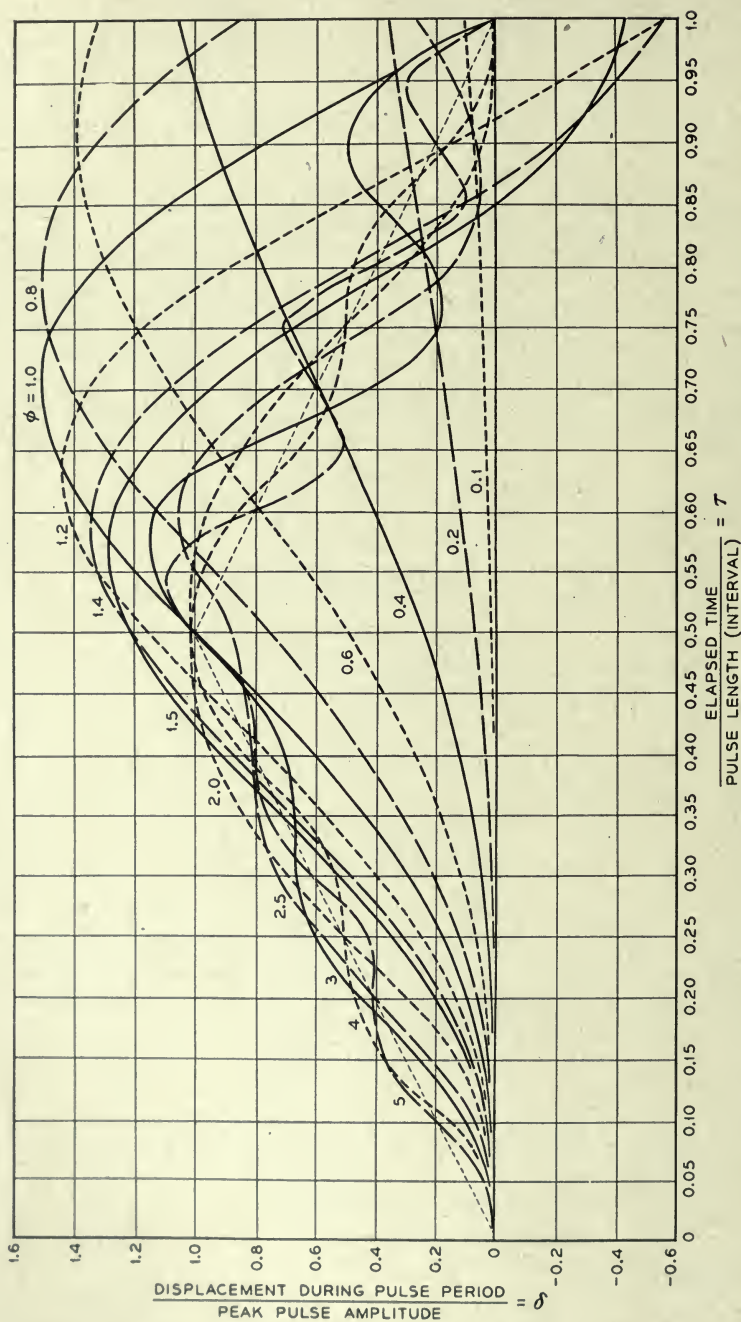


Fig. 4—Transient time displacement curves for various values of triangular whip.

From 1.04

$$F(s) = \frac{a\omega_0}{s^2 + \omega_0^2} (1 + e^{-\pi s/\omega_0}). \quad (13)$$

Substituting (13) in (5) we obtain

$$X(s) = \frac{a\omega^2\omega_0}{(s^2 + \omega^2)(s^2 + \omega_0^2)} (1 + e^{-\pi s/\omega_0}). \quad (14)$$

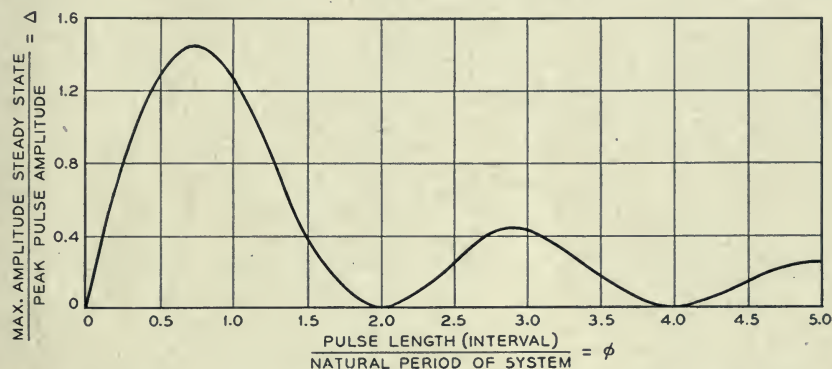


Fig. 5—Maximum amplitude as a function of frequency ratio—steady state triangular whip.

The inverse transform gives us

$$x = \frac{a\omega^2\omega_0}{\omega^2 - \omega_0^2} \left[\left(\frac{1}{\omega_0} \sin \omega_0 t - \frac{1}{\omega} \sin \omega t \right) + \left(\frac{1}{\omega_0} \sin \omega_0 \left(t - \frac{\pi}{\omega_0} \right) - \frac{1}{\omega} \sin \omega \left(t - \frac{\pi}{\omega_0} \right) \right) u \left(t - \frac{\pi}{\omega_0} \right) \right] \quad (15)$$

and dividing this into two parts again, the transient and the steady state, we find for the transient,

$$x = \frac{a\omega^2\omega_0}{\omega^2 - \omega_0^2} \left[\frac{1}{\omega_0} \sin \omega_0 t - \frac{1}{\omega} \sin \omega t \right]$$

and substituting the dimensionless quantities

$$\delta = \frac{x}{a}, \quad \frac{\pi/\omega_0}{T} = \varphi$$

and

$$\tau = \frac{t}{\pi/\omega_0}$$

we find

$$\delta = \frac{4\varphi^2}{4\varphi^2 - 1} \left(\sin \pi \tau - \frac{1}{2\varphi} \sin 2\pi \varphi \tau \right). \quad (16)$$

A plot of this equation as a family of curves for various values of φ is shown in Fig. 6. It is noted that, in general, this group of curves resembles those of Fig. 4 of the triangular whip.

The steady state is

$$x = \frac{a\omega^2\omega_0}{\omega^2 - \omega_0^2} \left[\frac{1}{\omega_0} \sin \omega_0 t - \frac{1}{\omega} \sin \omega t + \frac{1}{\omega_0} \sin \omega_0 \left(t - \frac{\pi}{\omega_0} \right) - \frac{1}{\omega} \sin \omega \left(t - \frac{\pi}{\omega_0} \right) \right] \quad (17)$$

and, in dimensionless quantities or expressed as a ratio of the pulse dimensions, we obtain

$$\delta_a = \frac{4\varphi}{1 - 4\varphi^2} \cos \pi\varphi \sin 2\pi\varphi(\tau - 1). \quad (18)$$

From (18) it follows that the maximum amplitude of the steady state is

$$\Delta = \frac{4\varphi \cos \pi\varphi}{1 - 4\varphi^2}. \quad (19)$$

A plot of this curve is shown in Fig. 7.

Shifted Cosine Whip.

The shifted cosine whip produces results of a similar nature. We have seen that the transform equation for this whip is (1.05)

$$F(s) = \frac{a\omega_0^2}{2s(s^2 + \omega_0^2)} (1 - e^{-2\pi s/\omega_0}). \quad (20)$$

Using equations (3) and (5) and transferring to dimensionless quantities, in which

$$\frac{x}{a} = \delta, \quad \frac{2\pi/\omega_0}{T} = \varphi = \frac{\omega}{\omega_0}, \quad \tau = \frac{t}{2\pi/\omega_0} = \frac{\omega_0 t}{2\pi}$$

we obtain

$$\delta = \frac{1}{2(\varphi^2 - 1)} \left(\cos 2\pi\varphi\tau - \varphi^2 \cos 2\pi\tau - \cos 2\pi\varphi(\tau - 1)u(\tau - 1) + \varphi^2 \cos 2\pi(\tau - 1)u(\tau - 1) \right). \quad (21)$$

Since we are interested only in the transient displacement, (21) becomes

$$\delta = \frac{(1 - \cos 2\pi\varphi\tau) - \varphi^2(1 - \cos 2\pi\tau)}{2(1 - \varphi^2)} \quad (22)$$

A family of curves showing δ for various values of φ is shown in Fig. 8.

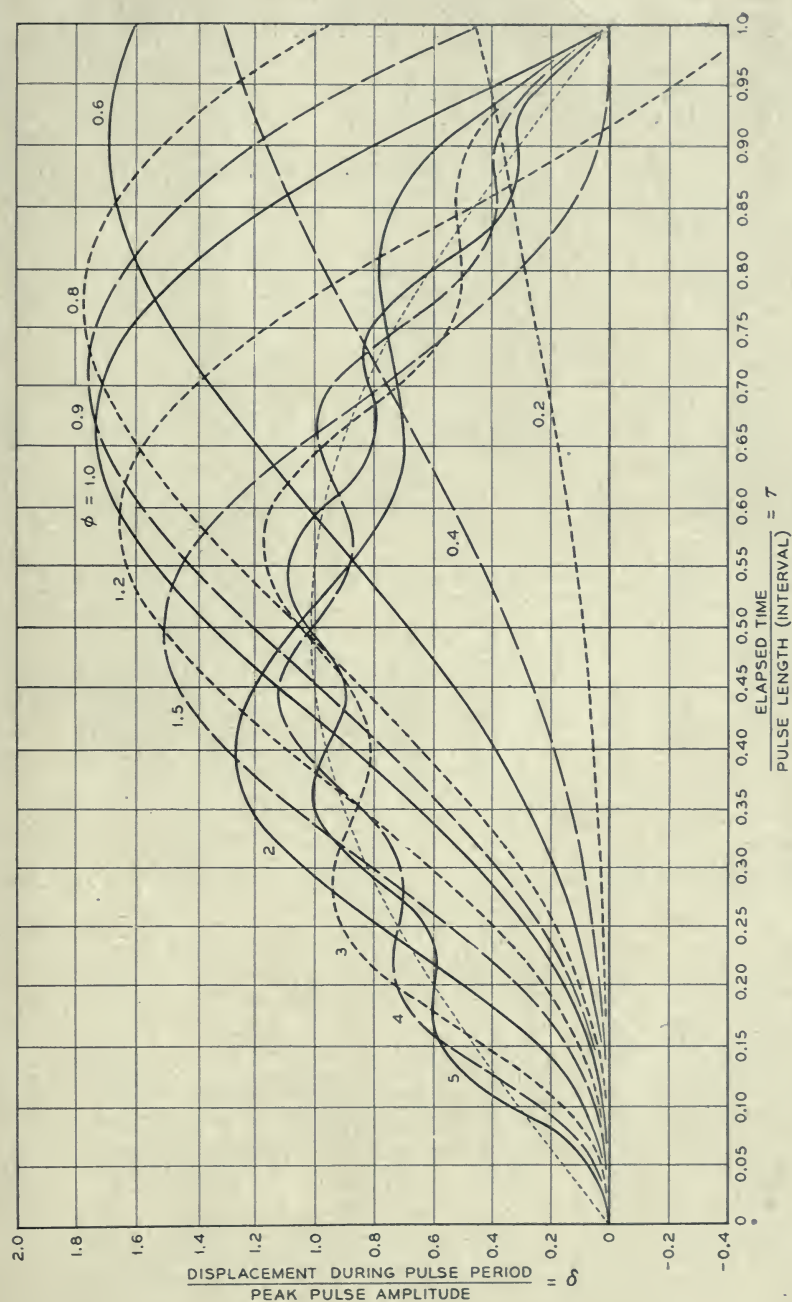


Fig. 6—Transient time displacement curves for various values of ϕ sine whip.

The steady state after the transient is (from Eq. 21)

$$\delta_a = \frac{1}{2(\varphi^2 - 1)} \left(\cos 2\pi\varphi\tau - \varphi^2 \cos 2\pi\tau - \cos 2\pi\varphi(\tau - 1) + \varphi^2 \cos 2\pi(\tau - 1) \right)$$

which reduces to

$$\delta_a = \frac{\sin \pi\varphi}{1 - \varphi^2} \sin (2\pi\varphi\tau - \pi\varphi). \quad (23)$$

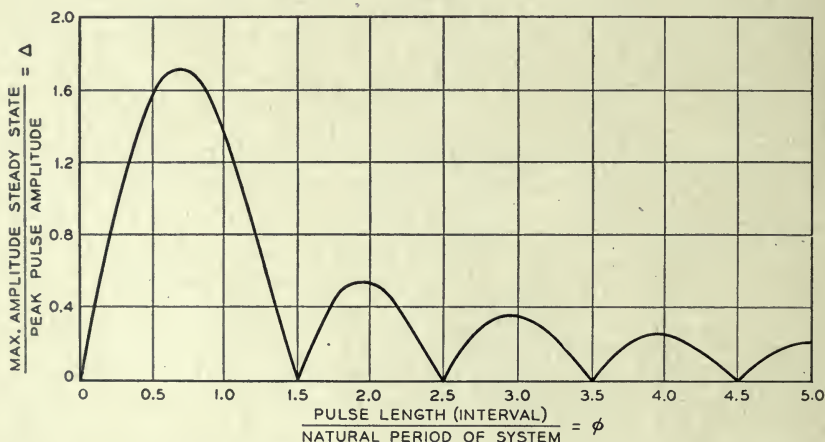


Fig. 7—Maximum amplitude as a function of frequency ration-steady state sine whip.

The maximum amplitude is

$$\Delta = \frac{\sin \pi\varphi}{1 - \varphi^2} \quad (24)$$

a plot of which is shown in Fig. 9.

Practical Considerations

Let us now consider the action of these various whips in terms of what they do to the system. The designer of shockmounts is primarily interested in the displacement across the mount or the relative displacement of base and mass.

In Fig. 10 the relative transient displacements for four systems are shown when subjected to a triangular whip. The natural frequencies are .4, 1.0, 1.5, and 2 times the frequency of the whip. From this it appears that the maximum relative displacement is approximately equal to the maximum

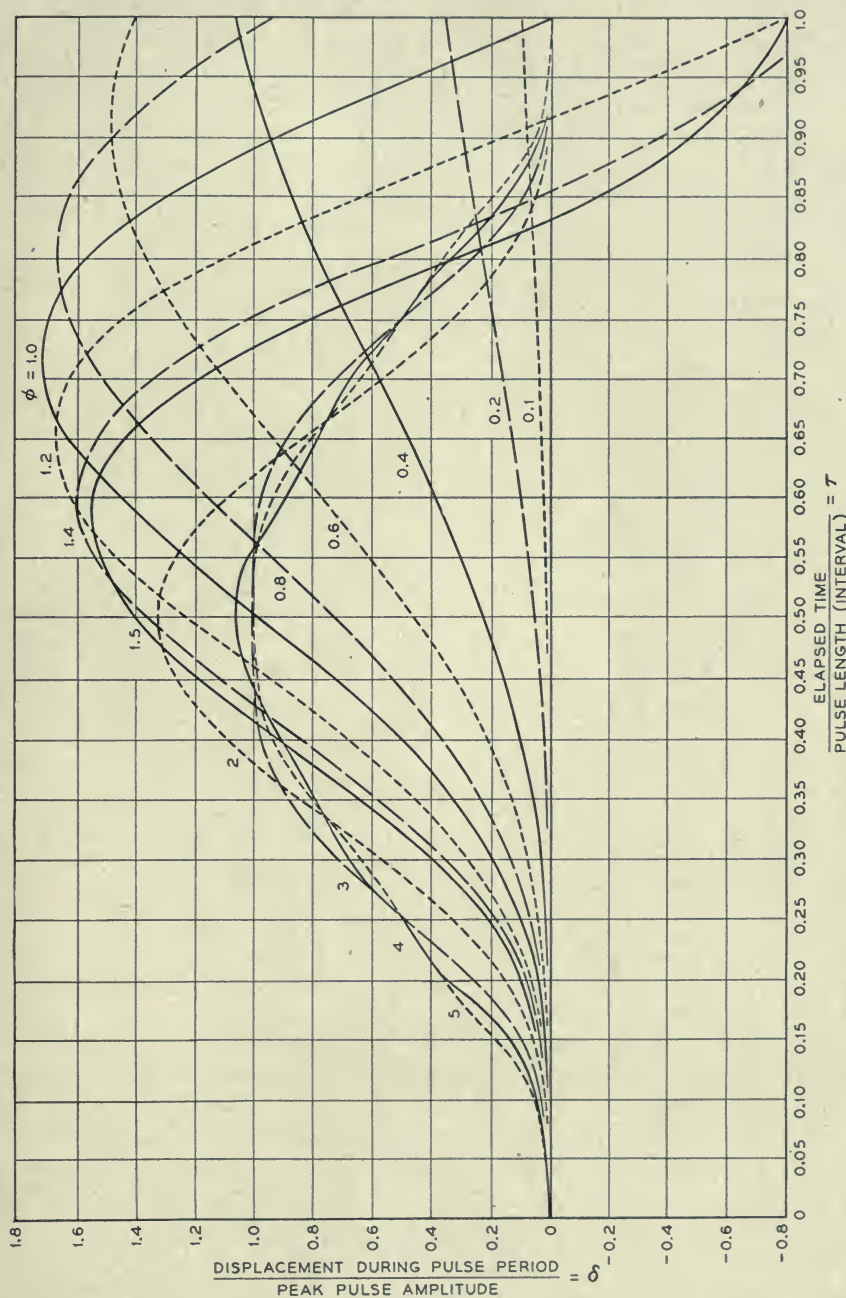


Fig. 8—Transient time displacement curves for various values of ϕ shifted cosine whip.

whip displacement. It is also observed that the large relative displacements occur when the frequency of the system is smaller than the pulse frequency.

After the transient has passed, the relative steady-state displacement, which is of course equal to the absolute, obtains large values too.

From Fig. 5 we note that a maximum of 1.5 is reached for the triangular whip and up to 1.7 times for the sine whip (see Fig. 7) at a frequency of approximately $\frac{3}{4}$ of the whip. Apparently even larger displacements across the mount occur after the transient has disappeared.

This is illustrated in Fig. 11 for the same systems as in Fig. 10.

As φ increases, which means if the frequency of the system increases with respect to the pulse frequency, the displacements across the mounts diminish,

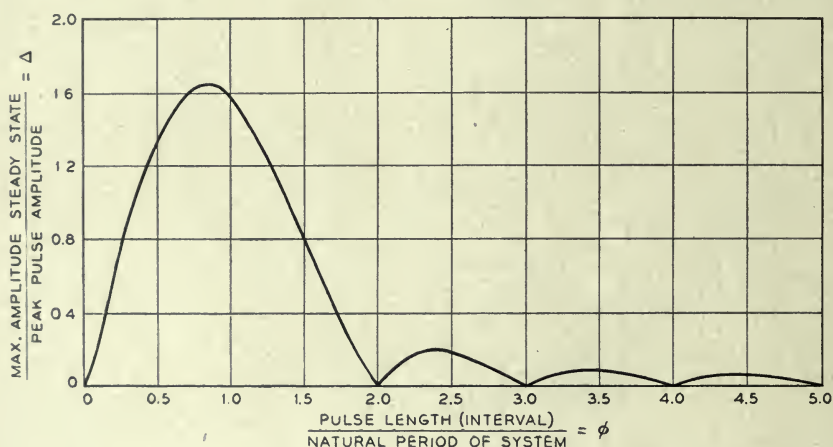


Fig. 9—Maximum amplitude as a function of frequency ratio—steady state shifted cosine whip.

while on the other hand the acceleration increases as will be shown later (see equation 39).

From this it seems advantageous to select a natural period of the system at least twice that of the pulse frequency.

The relative displacements are limited by practical considerations, such as available space between cabinets and bulk head, cable connections, personnel safety and others.

In the design of Bell Telephone Laboratories radar equipment, the relative displacement has been held to one-half inch, and the natural frequency in the neighborhood of 35 to 40 cycles per second or a period of 25 to 30 m.s.

The average of the heaviest shock administered to this type of equipment has a peak amplitude of 1.5 inch and a time interval of approximately 60 m.s.

From Fig. 5, we find that under these conditions a maximum relative dis-

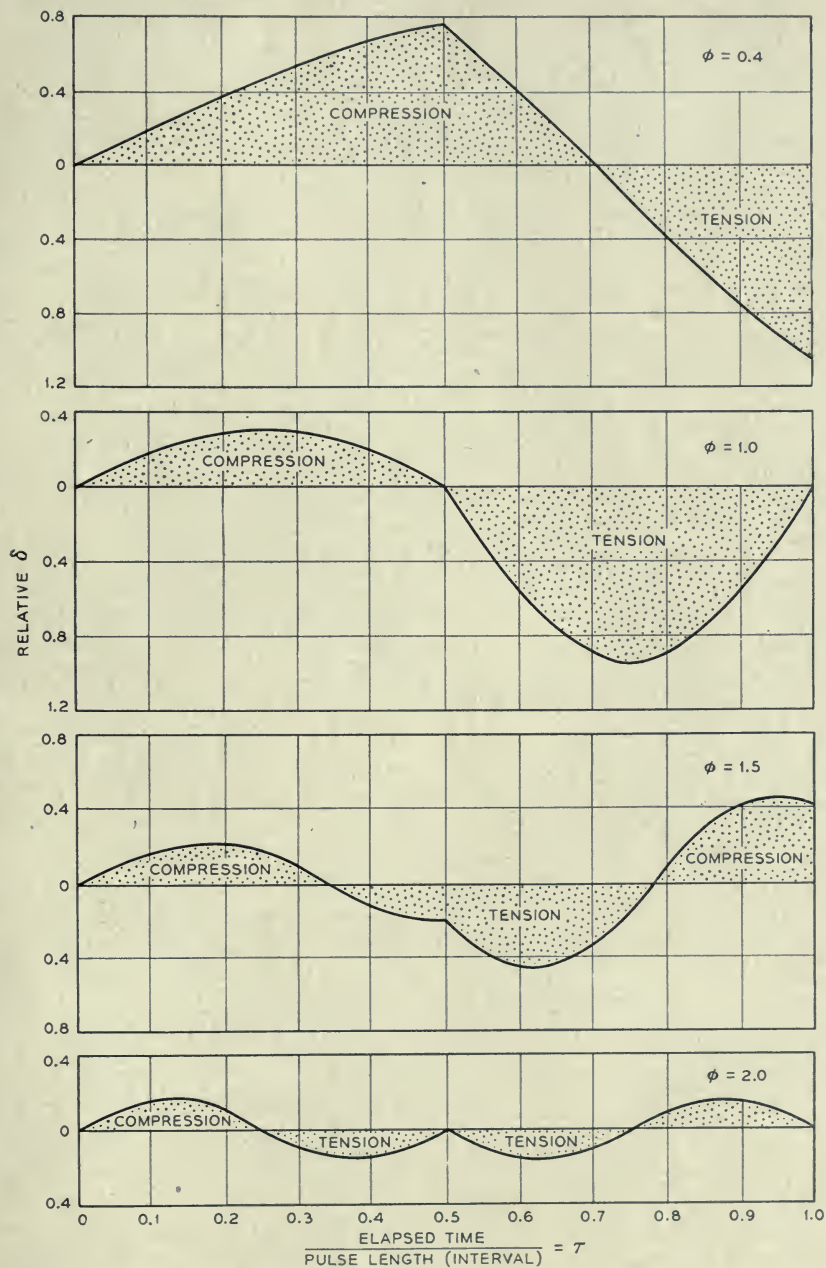


Fig. 10—Transient time displacement curves across the mount for various values of ϕ triangular whip.

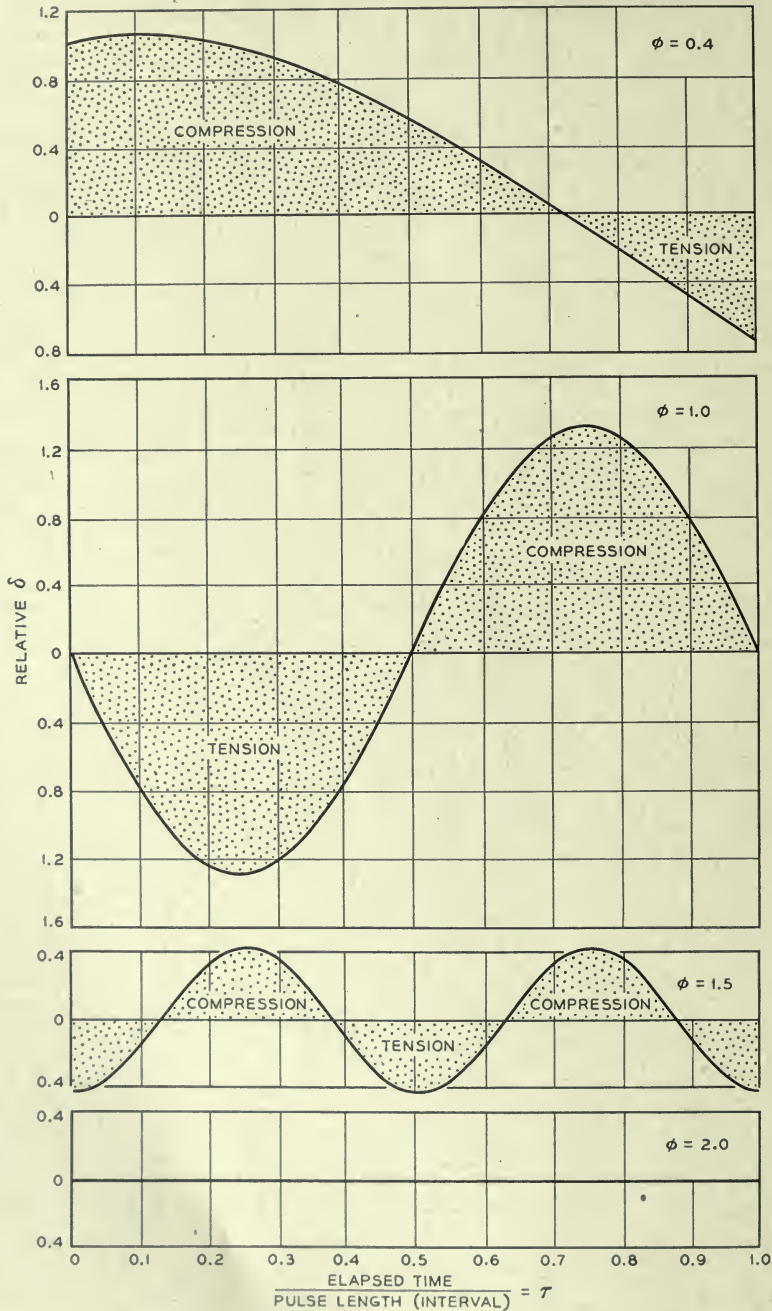


Fig. 11—Steady state time displacement curves across the mount. Triangular whip.

placement of .42 times the peak pulse amplitude or approximately $\frac{5}{8}$ inch may be expected.

Taking into consideration that the shock mount has been designed with a certain amount of damping, it is thus possible to hold the relative displacement within the boundaries of its shock-absorbing capacity.

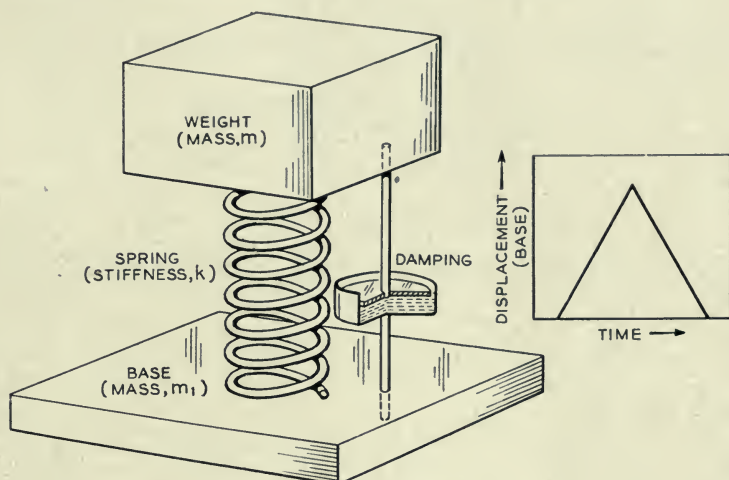


Fig. 12—System with damping.

Viscous Damping

The fundamental differential equation for a system with damping is (see Fig. 12).

$$\ddot{x} + 2\ell\dot{x} + \omega^2x = 0 \quad (25)$$

If we let a whip operate on this system we obtain

$$\ddot{x} + 2\ell\dot{x} + \omega^2x = \omega^2x_1(t) \quad (26)$$

However, the sudden displacement of the base also produces an acceleration of the mass proportional to the velocity. If $x_1(t)$ is the displacement then $\dot{x}_1(t)$ may be represented to be the velocity and $2\ell\dot{x}_1(t)$ the acceleration. We have, then, for the completed equation

$$\ddot{x} + 2\ell\dot{x} + \omega^2x = \omega^2x_1(t) + 2\ell\dot{x}_1(t) \quad (27)$$

In the Laplacian terminology, if

$$x_1(t) = F(s)$$

then

$$\dot{x}_1(t) = sF(s) \quad (\text{initial value being zero})$$

The Laplace transform of equation (27) is

$$X(s) = \frac{\omega^2 + 2\ell s}{s^2 + 2\ell s + \omega^2} F(s). \quad (28)$$

The solution of (28) is made easier if it is written in the form

$$X(s) = \frac{\omega^2 + 2\alpha s}{(s + \alpha)^2 + \beta^2} F(s) \quad (29)$$

in which

$$\alpha = \ell \quad \text{and} \quad \alpha^2 + \beta^2 = \omega^2.$$

Subjecting this system to a triangular whip, of which the Laplace transform is

$$F(s) = \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2$$

we have

$$X(s) = \frac{\omega^2 + 2\alpha s}{(s + \alpha)^2 + \beta^2} \frac{a}{b} \left(\frac{1 - e^{-bs}}{s} \right)^2 \quad (30)$$

the solution of which involves two transform pairs. The inverse transform gives us a solution of the transient as well as the steady state. It has been mentioned before that the steady state produces the maximum displacements across the mount; therefore it will be considered in more detail. We find that the steady state solution is

$$x_a(t) = \frac{a}{b} \frac{e^{-\alpha t}}{\beta} \left(-\sin \beta t + 2e^{\alpha b} \sin \beta(t - b) - e^{2\alpha b} \sin \beta(t - 2b) \right) \quad (31)$$

Which simplifies to

$$x_a(t) = \frac{e^{-\alpha t}}{\beta} \frac{a}{b} \sqrt{(A^2 + B^2)} \sin(\beta t - \theta) \quad (32)$$

Using dimensionless quantities

$$\eta = \frac{\ell}{\omega} = \frac{\alpha}{\omega} \quad \text{and} \quad \frac{x_a}{a} = \delta_a$$

and the substitution

$$\sqrt{1 - \eta^2} = \gamma,$$

we find that

$$b\beta = b\omega\gamma = \pi\phi\gamma$$

Equation (32) may now be expressed as

$$\delta_a = \frac{\epsilon^{-\alpha t}}{\pi\varphi} \sqrt{(A^2 + B^2)} \sin(\omega\gamma t - \theta) \quad \alpha t > 2\eta\pi\varphi \quad (33)$$

in which

$$\tan \theta = \frac{B}{A}$$

and

$$A = -1 + 2\epsilon^{\eta\pi\varphi} \cos \pi\varphi\gamma - \epsilon^{2\eta\pi\varphi} \cos 2\pi\varphi\gamma \quad (34)$$

$$B = 2\epsilon^{\eta\pi\varphi} \sin \pi\varphi\gamma - \epsilon^{2\eta\pi\varphi} \sin 2\pi\varphi\gamma \quad (35)$$

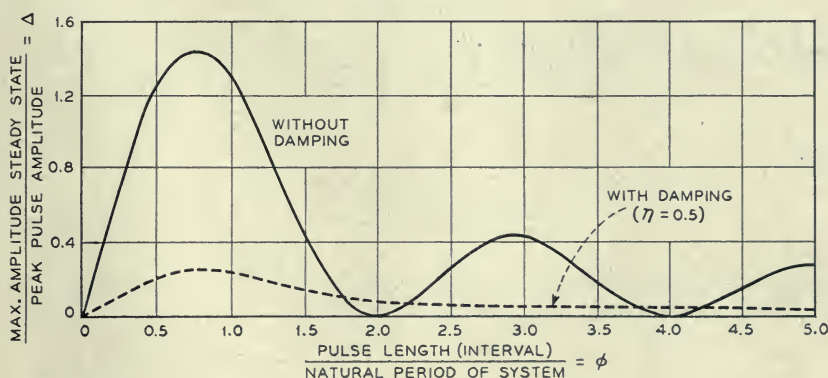


Fig. 13—Effect of damping on steady state amplitude for triangular whip.

From equation (33) we obtain the maximum displacement

$$\Delta = \frac{\epsilon^{-\alpha t}}{\pi\varphi} \sqrt{A^2 + B^2} \quad \alpha t > 2\eta\pi\varphi \quad (36)$$

in which

$$\alpha t = \frac{\eta}{\gamma} \left(\tan^{-1} \frac{\gamma}{\eta} + \tan^{-1} \frac{B}{A} \right) \quad \alpha t > 2\eta\pi\varphi$$

In Fig. 13 a plot of equation (36) is shown for $\eta = .5$. This indicates that the peak value of Δ is .24 as compared to 1.48 when no damping is present.

Accelerations

The transient accelerations of the mass m during the whip action and the subsequent steady state may be found by examining the acceleration during the first part of a triangular whip. Designating the velocity of displacement

of the whip by v the expression for x may be formed from equation (7) by setting $t < b$.

$$x = \frac{a\omega^2}{b} \left(\frac{t}{\omega^2} - \frac{\sin \omega t}{\omega^3} \right), \quad t < b \quad (37)$$

and, since $\frac{a}{b} = v$

$$x = vt - \frac{v}{\omega} \sin \omega t \quad (38)$$

whence

$$\ddot{x} = v\omega \sin \omega t$$

and the maximum acceleration is

$$A_0 = v\omega \quad (39)$$

Let

$$\frac{A_0}{v\omega} = \lambda_0 \text{ then } \lambda_0 = 1$$

By proceeding in a similar manner with the next step of the whip the value of the acceleration ratio will be found to be

$$\lambda_0 = 3$$

and for the completed whip or steady state

$$\lambda_0 = 4$$

The expression $A_0 = \lambda_0 v\omega$ is an important factor in shock considerations. Thus we have a simple relation for the final maximum amplitude of the periodic acceleration of the mass m when subjected to a triangular whip; viz., it is four times the product of whip velocity and natural frequency of the system. The constant λ_0 depends upon the configuration of the whip; the velocity v indicates the intensity of the whip; while ω expresses the kind of response the system is capable of.

It is of interest to note that this maximum periodic value of A_0 will be produced only if the ratio of pulse frequency and natural frequency is of the correct value. It is difficult to produce shocks on existing equipment of exactly the same characteristics within narrow limits as to time duration and therefore it must be expected that a considerable variation in damage may occur even though similar shocks are administered to identical test objects. For the same reason a shock of lower intensity may produce

more damage than a higher one because impulse amplitude as well as duration change at the same time.

Although damping is a highly desirable feature in a shock mount, the damping device may cause a certain amount of coupling between the mass and the base, and if too much damping is provided the transient acceleration of the mass may become excessive.

The analysis of this problem by means of the Laplace transforms is not difficult, for we can use results previously obtained. The transform equation for a system with damping, subjected to a whip, is

$$X(s) = \frac{\omega^2 + 2\alpha s}{(s + \alpha)^2 + \beta^2} F(s) \quad (40)$$

in which $F(s)$ represents the transform of the disturbance or excitation. Since we are interested in the effect of the damping or η upon the response, only the first part of the triangular whip will be considered.

In this case

$$x_1(t) = vt$$

and

$$\mathcal{L}[x_1(t)] = F(s) = \frac{v}{s^2}. \quad (41)$$

Substituting (41) in (40)

$$X(s) = \frac{v(\omega^2 + 2\alpha s)}{s^2[(s + \alpha)^2 + \beta^2]} = F_1(s) \quad (42)$$

If $X(s)$ is the transform of $x(t)$, a displacement, then the acceleration is $\ddot{x}(t)$ or $g(t)$, ($\dot{x}(t) = g(t)$ by definition) and

$$\mathcal{L}[\ddot{x}(t)] = \mathcal{L}[g(t)] = s^2 X(s)$$

Substitution in (42) gives

$$s^2 X(s) = 2v\alpha \frac{s + \frac{\omega^2}{2\alpha}}{(s + \alpha)^2 + \beta^2} \quad (43)$$

Now $\mathcal{L}^{-1}[s^2 X(s)] = g(t)$

so that

$$g(t) = \frac{2v\alpha}{\beta} \left[\left(\frac{\omega^2}{2\alpha} - \alpha \right)^2 + \beta^2 \right]^{\frac{1}{2}} e^{-\alpha t} \sin(\beta t + \psi) \quad (44)$$

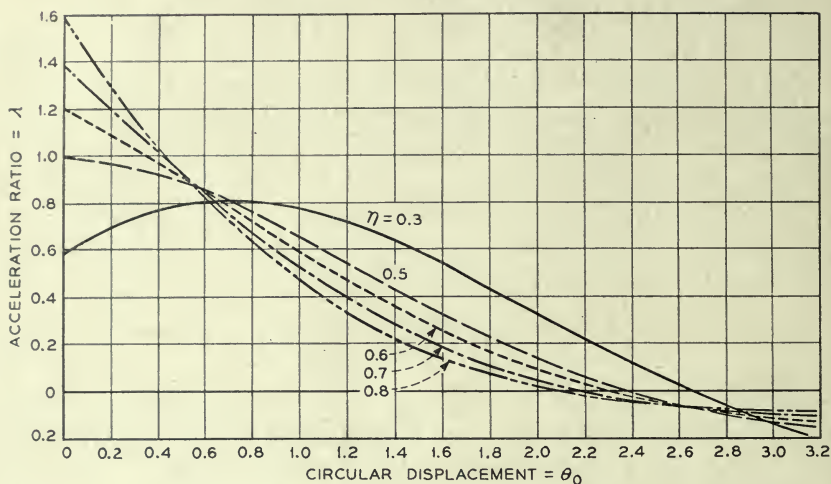


Fig. 14—Transient acceleration during initial part of triangular whip.

Letting

$$\frac{g(t)}{v\omega} = \lambda, \quad \frac{\alpha}{\omega} = \frac{\ell}{\omega} = \eta, \quad \beta = \omega \sqrt{1 - \eta^2} = \omega\gamma$$

and

$$\omega t = \theta_0$$

Substituting

$$\lambda = \frac{e^{-\eta\theta_0}}{\gamma} \sin(\lambda\theta_0 + \psi) \quad (45)$$

in which

$$\tan \psi = \frac{2\eta\gamma}{1 - 2\eta^2}$$

Figure 14 is a plot of λ against θ_0 for various values of η . It is noted that for $\eta = .5$ of critical damping the initial acceleration is equal to the undamped, λ being one.

The data presented here are also applicable to long duration pulses, because the final results have been given in dimensionless quantities, the units of measurement being those of the pulse.

SYMBOLS USED

Mass	m
Mass of base.....	m_1
Spring stiffness.....	k

Displacement mass.....	$x, x(t)$
Displacement base.....	$x_1, x_1(t)$
Force on base.....	$F, F(t)$
Velocity of base.....	v
Acceleration of base.....	$a_0, \ddot{x}_1$
Acceleration of mass m	$\ddot{x}, \ddot{x}(t), g(t)$
Maximum acceleration of the mass m	A_0
Maximum acceleration ratio.....	$\lambda_0 = \frac{A_0}{v\omega}$
Acceleration ratio.....	$\lambda = \frac{g(t)}{v\omega}$
Natural frequency of mass m (circular).....	ω, β_0
Circular (angular) displacement.....	$\omega t = \theta_0$
Frequency of sinusoid of which pulse consists (not pulse frequency).....	ω_0
Peak pulse displacement.....	a
Pulse period (triangular).....	$2b$
“ “ (sine pulse).....	$\frac{\pi}{\omega_0}$
“ “ (shifted cosine pulse).....	$\frac{2\pi}{\omega_0}$
Period of mass m	T
$\frac{\text{displacement during pulse period}}{\text{peak pulse amplitude}} =$	$\frac{x}{a} = \delta$
$\frac{\text{amplitude steady state}}{\text{peak pulse amplitude}} =$	δ_a
$\frac{\text{max. amplitude steady state}}{\text{peak pulse amplitude}} =$	Δ
$\frac{\text{elapsed time}}{\text{pulse length (interval)}} =$	$\frac{t}{2b}, \frac{t/\pi}{\omega_0}, \frac{t/2\pi}{\omega_0}, \frac{t}{T_0} = \tau$
$\frac{\text{pulse length (interval)}}{\text{natural period of system}} =$	$\frac{T_0}{T} = \varphi$
Damping coefficient.....	ℓ, α
Critical damping ratio.....	$\frac{\ell}{\omega} = \eta$
Transform of $x(t) =$	$X(s)$
“ “ $x_1(t) =$	$X_1(s)$
“ “ $F(t) =$	$F_0(s)$
“ “ $f(t) =$	$F(s)$
$f(t)$ represents any function of t , without reference to its dimensional magnitude.	

Maximally-flat Filters in Waveguide

By W. W. MUMFORD

Microwave radio relay repeaters require the use of band-pass filters which match closely the impedances of the interconnecting transmission lines and which suppress adjacent channels adequately. A type of structure called a *Maximally-Flat* filter meets these requirements.

The ladder network which gives a maximally-flat insertion loss characteristic is discussed and several methods of achieving its counterpart in microwave transmission lines are presented. Resonant cavities are used to simulate tuned circuits and the necessary formulas relative to this approximate equivalence are given.

Experimental data confirm the theory and show that this technique yields remarkable impedance matches.

INTRODUCTION

WE USUALLY associated the word *filter* with any device which is selective. The electric wave filter has that property which enables it to transmit energy in one band or bands of frequencies and to inhibit energy in other bands. Selectivity is the result of either selective absorption^{1, 2*} or selective reflection. This paper discusses a special case of the classical lossless transducer which derives its selective properties entirely from selective reflection. The insertion loss of this type of filter can be analyzed in terms of the input reflection coefficient and the input standing wave ratio.

In many applications of lossless filters it is desirable to obtain a characteristic such that the insertion loss, and hence the reflection coefficient, is small over as wide a band as possible. A special case described here, referred to as a maximally-flat filter, has a loss characteristic such that a maximum number of its derivatives are zero at midband. While the maximally-flat type of characteristic does not give the smallest possible reflections over a finite pass band, it does give small reflections, and has added advantages of simplicity in design and in many cases less transient distortion than filters giving smaller reflections.

The desirable characteristics of maximally-flat filters have long been realized.^{1, 3} Mr. W. R. Bennett⁴ of these Laboratories derived the constants for a maximally-flat ladder network in the late 20's, and gave simple expressions for the element values. Butterworth,⁵ Landon⁶ and Wallman⁷ have treated maximally-flat filter-amplifiers in which the filter sections are separated by amplifier tubes. Darlington⁸ has considered the general case of four terminal filters which have insertion loss characteristics that can be prescribed, but he places the emphasis more on filters that have tolerable

* A list of selected references appears at the end of the paper.

ripples in the pass-band than on maximally-flat structures. The work of Bennett will be followed closely not only because it came first, but also because it is easy to understand.

Bennett expresses the values of the filter branches in terms of their cutoff frequencies, which in turn bear a relationship to the cutoff frequencies of the total filter. In the language of one who is familiar with microwave technique,^{9, 10, 11, 12} the values of the filter branches can be expressed in terms of the loaded Q 's of the cavities, which in turn bear a relationship to the loaded Q of the total filter. A simple mathematical expression connects the loaded Q to the cutoff wavelengths.

At low frequencies the band-pass maximally-flat filter is made up of resonant branches connected alternately in series and in parallel. The microwave analogue of this configuration is obtained by the use of shunt resonant cavities that are spaced approximately a quarter wavelength apart in the waveguide. Use is made of the impedance inverting property of a quarter wave line, thereby eliminating the necessity of using both series and parallel branches.

The resonant cavity in the waveguide resembles a shunt resonant tuned circuit,¹³ but is different in several minor respects. An analysis of these differences reveals the corrective measures that are necessary in order that the simulation shall be sufficiently accurately attained.

The first part of the paper deals with the concepts of loaded Q and resonant filter branches of both the series and the parallel types. Admittance and impedance functions, as well as expressions for the insertion loss, are given using these terms, and the relationship between loaded Q and cutoff frequencies is stated. This concept of loaded Q is then introduced to describe the performance of a complete maximally-flat filter in terms of its cutoff frequencies. The insertion loss is then given as a simple expression containing the total Q and the resonant frequency. The Q 's of all the branches are derived from the total Q in simple terms. The connection between the insertion loss and the input standing wave ratio is then discussed before turning to the actual design problem.

Next the paper deals with the application of the filter theory to waveguide technique. The limitations of the quarter-wave coupling lines are pointed out and the added selectivity due to them is derived.

Then the paper compares microwave resonant cavities with parallel-tuned circuits. Formulas are given which relate the geometrical configuration to the loaded Q , the resonant frequency and the excess phase of the cavities. Three types of cavities are treated: those using inductive posts, inductive irises and capacitive irises.

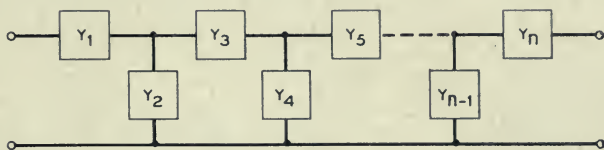
Finally, the measured results on a four-cavity maximally-flat filter in $1'' \times 2''$ waveguide are presented and compared with the original design points.

As a further confirmation of the theory, the experimental results on another and longer waveguide filter consisting of fifteen resonant cavities and fourteen connecting lines are given. The conclusion is reached that maximally-flat waveguide filters can be designed to have excellent impedance match and off-band suppression qualities.

NOTATION

a	Width of waveguide.
b	Height of waveguide.
B	Normalized susceptance.
c	Velocity of light in free space.
C_r	Capacitance in the r^{th} branch of a filter.
d	Width of iris opening.
d	Diameter of post in waveguide.
δ	A small number $\ll 1$.
e	Base of natural logarithms.
f	Frequency.
f_0	Resonant frequency.
f_c	Frequency at half power point.
f_{cw}	Cutoff frequency of waveguide.
G	Terminating conductance of filter.
θ	$\frac{2\pi\ell}{\lambda}$
K	Susceptance parameter.
ℓ	Length of transmission line.
ℓ'	Length of line corresponding to excess phase of cavity.
ℓ_c	Length of line connecting two cavities.
λ_0	Resonant wavelength.
λ_c	Wavelength at half power point.
λ_g	Wavelength in transmission line.
λ_a	Wavelength in free space.
L_r	Inductance in r^{th} branch of a filter.
m	An integer, including zero.
n	Number of branches in filter.
P_0	Available power.
P_L	Power delivered to load.
Q	Loaded Q . The selectivity of a loaded circuit.
Q_r	Loaded Q of the r^{th} branch.
Q_T	Loaded Q of the total filter.
R	Terminating resistance of the filter.
s	Distance from center of waveguide.
S	Voltage standing wave ratio.

τ	Thickness of iris.
V_{max}	Maximum voltage on transmission line.
V_{min}	Minimum voltage on transmission line.
ω	Angular frequency.
Ω	Frequency parameter.
Y	Admittance.
Y_0	Surge admittance of transmission line.
Z	Impedance.
Z_0	Surge impedance of transmission line.



THE Y'S ARE USED TO DENOTE
GENERALIZED ADMITTANCE FUNCTIONS

Fig. 1—Block diagram of a filter consisting of a ladder network.

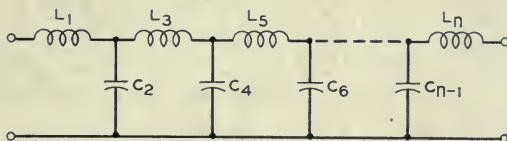


Fig. 2—Schematic diagram of a low-pass filter.

GENERAL

The art of designing filters which utilize lumped elements is well known. Desirable characteristics may be obtained by means of a ladder network of generalized admittances, such as is illustrated in Fig. 1. In particular a low-pass filter takes the configuration shown in Fig. 2, and a band-pass filter that of Fig. 3. In either case, certain frequency selectivity characteristics can be obtained when the individual branches are assigned definite values. The individual branches, L_1C_1 , L_2C_2 , in the bandpass filter consist of an inductance, L_r , and capacity, C_r , in series or shunt. For the specific case to be discussed in this paper, namely a filter consisting of lossless elements intended for insertion between a source having an internal resistance R and a receiver having the same resistance, analysis is simplified if a branch is described in terms of its resonant frequency and its loaded Q .*

* The loaded Q of a resonant branch in such a filter is the reciprocal of its percentage band width measured to the half power points when that branch alone is fed by the same generator and has the same load resistance as that of the total filter.

resonant frequency of a branch is independent of the terminal resistance and is given by the relation

$$f_0 = \frac{1}{2\pi\sqrt{L_r C_r}} \quad (1)$$

The loaded Q of the branch, to be designated Q_r , is a function not only of the inductance and capacitance in the branch, but also of the resistance, R , of the terminations on the filter. For the series resonant branches, it is given by

$$Q_r = \frac{1}{2R} \sqrt{\frac{L_r}{C_r}} = \frac{\omega_0 L_r}{2R} \quad (2)$$

and for the parallel resonant branches

$$Q_r = \frac{R}{2} \sqrt{\frac{C_r}{L_r}} = \frac{\omega_0 C_r}{2} R. \quad (3)$$

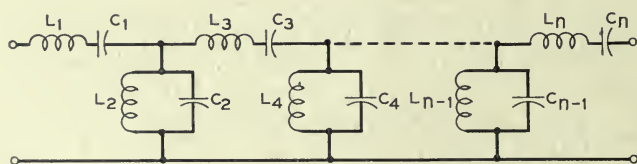


Fig. 3—Schematic diagram of a band-pass filter.

It may be noted that the loaded Q can be defined in terms of the insertion loss imposed by connecting the branch between a source and receiver each of resistance R . Analysis of such a circuit shows that

$$\frac{P_0}{P_L} = 1 + Q_r^2 \left(\frac{f}{f_0} - \frac{f_0}{f} \right)^2 \quad (4)$$

where f is the frequency;

P_0 is the power available from a generator which has an internal resistance R ;

P_L is the power delivered through the inserted branch to a load of resistance R .

At the cutoff frequency, f_c , defined as the frequency at which the power delivered to the load is half the available power, $\frac{P_0}{P_L} = 2$, whence

$$Q_r = \left| \frac{1}{\frac{f_c}{f_0} - \frac{f_0}{f_c}} \right| = \left| \frac{f_0}{f_{c2} - f_{c1}} \right|. \quad (5)$$

Written in terms of the wavelengths this becomes

$$Q_r = \frac{\frac{1}{\lambda_0}}{\frac{1}{\lambda_{c_2}} - \frac{1}{\lambda_{c_1}}} \quad (6)$$

This equation is a convenient one to use later in the discussion on resonant cavities.

The normalized admittance of a single-shunt branch terminated by a resistance R can be expressed in terms of its resonant frequency and its Q ; thus

$$YR = 1 + j2Q_r \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \quad (7)$$

Similarly, the normalized impedance of a single-series branch terminated by a resistance R can be written

$$\frac{Z}{R} = 1 + j2Q_r \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \quad (8)$$

The use of the term loaded Q thus has the advantage that expressions for normalized admittance and normalized impedance of shunt and series resonant circuits respectively are identical, as are also the corresponding expressions for their insertion loss functions.

Loss functions of complete filters can likewise be expressed in terms of a loaded Q defined for the complete filter. For example, the loss function of the particular type of filter called a "Maximally-flat" filter is given⁴

$$\frac{P_0}{P_L} = 1 + \left[\frac{\frac{f}{f_0} - \frac{f_0}{f}}{\frac{f_c}{f_0} - \frac{f_0}{f_c}} \right]^{2n} \quad (9)$$

where n is the number of resonant branches in the filter, and f_c is the cutoff frequency of the filter (half power points).

In consequence of the concept of loaded Q of the total filter, the loss function can be expressed as

$$\frac{P_0}{P_L} = 1 + \left[Q_r \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \right]^{2n} \quad (10)$$

where the total Q_r of the filter is

$$Q_r = \left| \frac{1}{\frac{f_c}{f_0} - \frac{f_0}{f_c}} \right| \quad (11)$$

For convenience, the bracketed term of equation 9 may be called Ω , a frequency parameter, whence

$$\Omega = \left[\frac{\frac{f}{f_0} - \frac{f_0}{f}}{\frac{f_c}{f_0} - \frac{f_0}{f_c}} \right] = Q_r \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \quad (12)$$

and the loss function becomes

$$\frac{P_0}{P_L} = 1 + (\Omega)^{2n}. \quad (13)$$

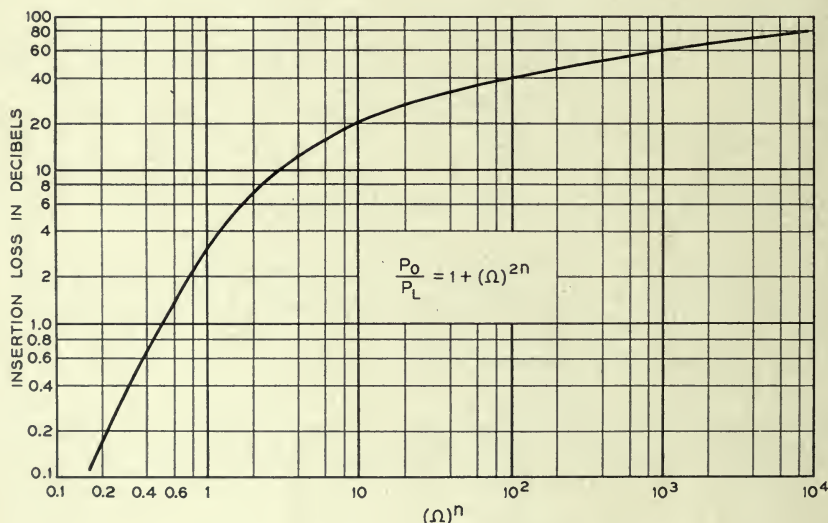


Fig. 4—Insertion loss of maximally-flat filters.

MAXIMALLY-FLAT FILTERS

The loss function for maximally-flat filters as given in equation 13 is plotted in Fig. 4 where the insertion loss in db is used on the ordinate and Ω^n is used on the abscissa.

The ladder network which gives rise to this loss function consists of n resonant branches, as shown in Fig. 3, that are all tuned to the same frequency, but whose selectivities, or loaded Q 's, are tapered from one end of the filter to the other according to the positive imaginary parts of the $2n$ roots of -1 , according to the theories of Bennett⁴ and Darlington.⁸ These roots are expressed thus

$$\sin \left(\frac{2r-1}{2n} \right) \pi$$

where r is the number of the root, n is the total number of branches. Thus the selectivities of the branches follow the relation

$$Q_r = Q_T \sin \left(\frac{2r-1}{2n} \right) \pi \quad (14)$$

where Q_T represents the selectivity of the total filter, and Q_r represents the required selectivity of the r^{th} branch, e.g., the selectivities of the first, second and third branches are

$$\begin{aligned} Q_1 &= Q_T \sin \frac{\pi}{2n} \\ Q_2 &= Q_T \sin \frac{3\pi}{2n} \\ Q_3 &= Q_T \sin \frac{5\pi}{2n}. \end{aligned} \quad (15)$$

This type of filter is particularly practical when a filter is required to give more than a certain amount of insertion loss in an adjacent band, and less than another certain amount of insertion loss at the edges of the pass-band. Putting this information in equation 10 gives two equations containing two unknowns, Q_T , the selectivity of the total filter, and n , the number of branches needed to fulfill the stated requirements. The solution for n may be fractional, in which event the next higher integral value of n is chosen, and this value is used to determine the selectivity, Q_T , of the filter. From this, the selectivities of all the branches are determined in accordance with equation 14.

STANDING WAVE RATIO

An alternative way of specifying filter performance is to refer to the input impedance mismatch as a function of frequency. The impedance mismatch can be expressed in terms of the direct and the reflected waves and in terms of the standing wave ratio that exists along the transmission line that connects the properly terminated filter with its generator. The standing wave ratio and the insertion loss of a filter bear a definite relationship to each other if the filter is composed of purely reactive elements. This relationship is given by the formula

$$\frac{P_0}{P_L} = \frac{(S+1)^2}{4S} \quad (16)$$

where S is the standing wave ratio, $\frac{V_{\max}}{V_{\min}}$, of the maximum voltage to the minimum voltage as measured along the transmission line.

When the filter characteristic is given by equation 13, the relationship between Ω , the frequency parameter, and the standing wave ratio can be expressed as

$$(\Omega)^n = \frac{S - 1}{2\sqrt{S}}. \quad (17)$$

This is shown graphically in Fig. 5, where the standing wave ratio is given in db $\left(20 \log_{10} \frac{V_{\max}}{V_{\min}}\right)$. This graph is used as an aid in the design of filters of

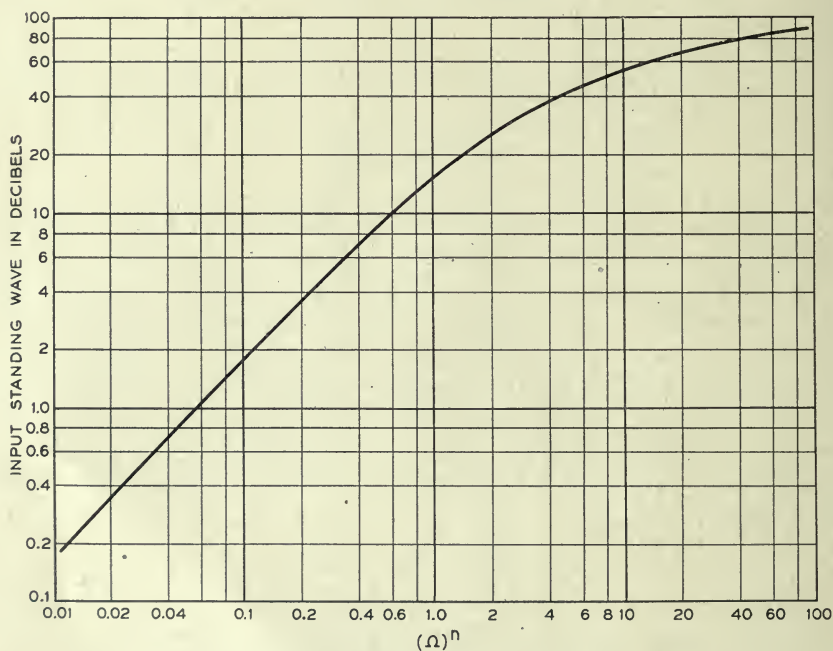


Fig. 5—Input standing wave ratio of maximally-flat filters.

this type, where the requirements are given in terms of the standing wave ratio. From this information the number of filter branches and the selectivity of the total filter can be determined, either from equation 17 or from Fig. 5.

DISTRIBUTED BRANCHES

It has been assumed that the mutual impedances of successive branches are all zero. At low frequencies this limitation may not be a serious one and the practical realization of the expected filter characteristics is accomplished by shielding properly one branch from another. However, as the

frequency is increased it becomes difficult to isolate the branches and undesirable mutual impedances arise which complicate the problem. In particular, in the microwave region, where waveguides are used, the physical size of each branch may be large compared with the wavelength and it is then impossible to lump all the branches at one place in the waveguide without encountering the complicated effect of mutual impedances.

A practical way of circumventing this difficulty is to distribute the branch circuits along the transmission line or waveguide at such distances that the mutual impedances become negligible. Then, however, the lengths of transmission line act as transducers, but since their properties are well understood and readily calculable this appears to be a practical solution. As a matter of fact, the impedance transforming properties of a length of transmission line can be used to advantage.^{9, 10, 14} For instance, it is well known that a quarter wavelength of lossless line transforms a load impedance according to the relation

$$Z = \frac{Z_0^2}{Z_L} \quad (18)$$

where Z_0 is the surge impedance of the line and Z_L is the load impedance.

Hence if the load impedance consists of a series resonant circuit containing an inductance, a capacity and a resistance equal to Z_0 in series, the impedance at the input end of the quarter wavelength of line is given

$$Z = \frac{Z_0}{\left[1 + j2Q \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \right]}. \quad (19)$$

The input admittance is

$$Y = Y_0 \left[1 + j2Q \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \right]. \quad (20)$$

As can be seen from equation 7, this is identical with the input admittance of a parallel tuned circuit whose terminating conductance is

$$G = Y_0. \quad (21)$$

The quarter-wave line likewise transforms a parallel circuit to a series circuit, as is illustrated in Fig. 6. This property of the quarter-wave line thus makes it possible to simulate a ladder network of alternate series and shunt branches by spacing shunt branches (or series branches) at quarter wavelength intervals along a transmission line, as illustrated in Fig. 7. The resonant frequencies and the selectivities of the branches are chosen as before.

Sometimes in practice a quarter wavelength may not be sufficient spacing to avoid mutual impedances arising between adjacent elements, in which event the connecting line may be increased to a higher odd multiple of quarter wavelength. This accentuates the frequency sensitivity of the connect-

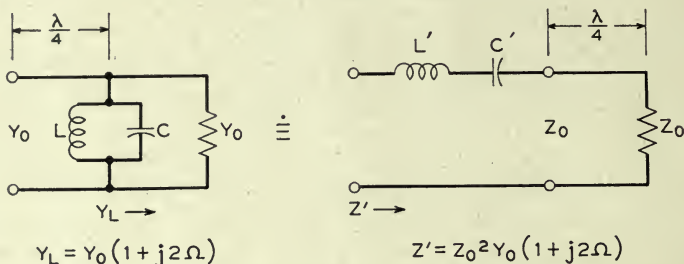
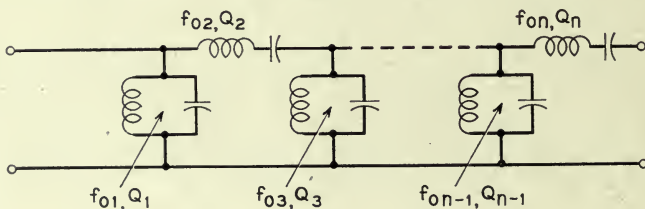


Fig. 6—Illustrating the impedance inverting property of a quarter wavelength of transmission line.

LUMPED CONSTANT FILTER USING SERIES & SHUNT ELEMENTS



LUMPED CONSTANT FILTER USING ONLY SHUNT ELEMENTS

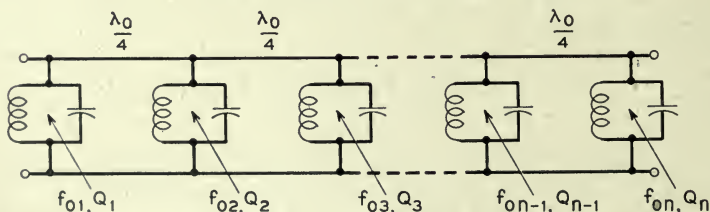


Fig. 7—Simulation of ladder network by shunt branches at quarter wave intervals.

ing line, but this effect can be taken into account by decreasing the selectivities of the branches themselves by appropriate amounts. In narrow-band filters this may be negligible, but in broad-band filters it may be considerable, as shown in the following analysis.

SELECTIVITY OF CONNECTING LINES

Consider a length of transmission line having a surge impedance $Z_0 = \frac{1}{Y_0}$ and terminated in a parallel resonant circuit containing an inductance, a

capacitance and a resistance equal to Z_0 ,* as in Fig. 6. The terminating admittance is given by the relation (See Eq. 7 and 9)

$$Y_L = Y_0(1 + j2\Omega). \quad (22)$$

The input admittance at the end of the length of line, ℓ , (nominally a quarter wavelength long) is given by the relation

$$\frac{Y}{Y_0} = \frac{(1 + j2\Omega) \cos \theta + j \sin \theta}{\cos \theta + j(1 + j2\Omega) \sin \theta} \quad (23)$$

where

$$\theta = \frac{2\pi\ell}{\lambda}$$

ℓ = length of line

λ = wavelength

$$\Omega = Q \left(\frac{f}{f_0} - \frac{f_0}{f} \right)$$

letting

$$\theta = \frac{\pi}{2} (1 + \delta) = \frac{\pi}{2} + \frac{\pi\delta}{2} \quad (24)$$

$$\cos \theta = -\sin \frac{\pi\delta}{2} \doteq -\frac{\pi\delta}{2} \quad (25)$$

$$\sin \theta = \cos \frac{\pi\delta}{2} \doteq 1 \quad (26)$$

where δ is a number small compared with 1. Then the admittance becomes

$$\frac{Y}{Y_0} \doteq j \frac{\pi\delta}{2} + \frac{1}{1 + j \left(2\Omega + \frac{\pi\delta}{2} \right)}. \quad (27)$$

This is the normalized input admittance of a circuit as shown in Fig. 8, where each end of an ideally inverting line is shunted by a tuned circuit whose normalized admittance is $j \frac{\pi\delta}{2}$.

From Eq. 24, setting $\frac{2\pi\ell}{\lambda_0} = \frac{\pi}{2}$, it follows that

$$\delta = \pm \left(\frac{f}{f_0} - 1 \right) \doteq \left(\frac{f}{f_0} - \frac{f_0}{f} \right) \left(\frac{1}{2} \right). \quad (28)$$

* More generally, the terminating admittance can assume any value without affecting the result.

From equation 7, the admittance of the circuit is expressed in terms of its selectivity, thus

$$j \frac{\pi \delta}{2} = j2Q \left(\frac{f}{f_0} - \frac{f_0}{f} \right). \quad (29)$$

Solving for the selectivity of this circuit, from Equations 28 and 29:

$$Q = \frac{\pi}{8}. \quad (30)$$

The selectivity of the coupling line can hence be counteracted by subtracting $\frac{\pi}{8}$ from the selectivities of the branches associated with it, provided

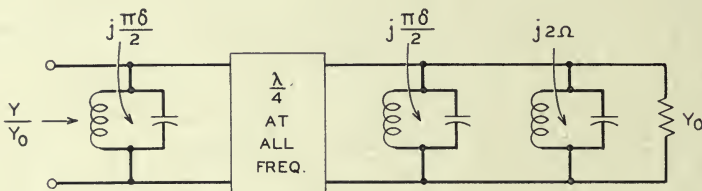


Fig. 8—Schematic diagram illustrating that the selectivity of a quarter wavelength of line can be represented by adding a tuned circuit to each end of an ideally inverting impedance transformer.

the coupling line is a quarter wavelength long. If it becomes necessary to use $\frac{3}{4}$ wavelength coupling lines, the selectivity of the line is tripled and $\frac{3\pi}{8}$ is subtracted from the selectivities of the associated branches.

RESONANT CAVITIES

The foregoing analysis reviews the principles of the design of filters which use lumped-constant circuits distributed along a transmission line. These principles can be applied to the design of filters in waveguides, coaxial lines, or any other types of transmission lines, provided that these lines are sufficiently lossless, the band is sufficiently narrow and the branches themselves are realizable. In the microwave region the first two provisions are usually met without difficulty, as is also the third provision when circuits with distributed constants are used. It may be difficult to construct a coil and a condenser circuit for microwaves, but easy to construct a resonant cavity which displays some of the desirable properties of the tuned circuit. Resonant cavities are similar to lumped tuned circuits in two respects.^{12, 13} They transmit a band of frequencies and they introduce a phase shift. An approximate equivalence is demonstrated in Appendix I, and is illustrated in Fig. 9, which depicts a resonant cavity as being nearly identical with a

tuned circuit situated across the middle of a short length of transmission line. This short length of transmission line is added in order to account for an excess of phase shift associated with the resonant cavity, but it can readily be absorbed in the connecting line which otherwise would have been an odd quarter wavelength long.

The similarity between resonant cavities and resonant lumped circuits enables one to use the known art of designing lumped element filters to de-

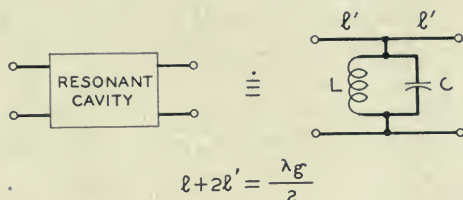
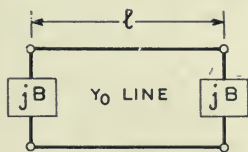


Fig. 9—A resonant cavity is approximately equivalent to a resonant circuit shunted across a short length of transmission line.



$$\tan \frac{2\pi\ell}{\lambda_0} = \frac{2}{B}$$

$$Q = \frac{\text{ARC TAN } \frac{2}{B}}{2 \text{ ARC SIN } \frac{2}{\sqrt{B^4 + 4B^2}}} = \frac{\sqrt{B^4 + 4B^2}}{4} \text{ ARC TAN } \frac{2}{B}$$

Fig. 10—The resonant wavelength and the loaded Q of a cavity depend upon the normalized susceptance of the end obstacles and their separation.

sign filters which use resonant cavities, provided that the selectivity, the resonant frequency and the excess phase shift of the resonant cavity are known.

RESONANT WAVELENGTH AND LOADED Q OF CAVITIES

These properties can best be derived by considering one of the usual types of cavities, which consists of two obstacles or discontinuities separated by a length of transmission line. Such a cavity is shown schematically in Fig. 10. The obstacles at each end are assumed to be equal, and to have an unvarying susceptance BY_0 , where Y_0 is the surge admittance of the con-

necting transmission line. This type of cavity is resonant when the relation is satisfied^{9, 10}

$$\tan \frac{2\pi\ell}{\lambda_0} = \frac{2}{B} \quad (31)$$

where λ_0 is the resonant wavelength in the transmission line,

ℓ is the length of the cavity

B is the normalized susceptance of the end obstacles.

This resonance occurs at any number of wavelengths, but the 1st or 2nd longest wavelength at which resonance occurs is in the region which is usually of greatest interest.

The selectivity in this region is determined also by the value of the normalized susceptance, B , of the obstacles, and is given by the relation (See Appendix I)

$$Q = \frac{\arctan \frac{2}{B}}{2 \arcsin \frac{2}{\sqrt{B^4 + 4B^2}}} \quad (32)$$

This selectivity is based upon the wavelength, not the frequency parameter. In terms of the wavelength in the transmission line this is

$$Q = \left| \frac{\frac{2\pi\ell}{\lambda_{g0}}}{\frac{2\pi\ell}{\lambda_{gc1}} - \frac{2\pi\ell}{\lambda_{gc2}}} \right| = \left| \frac{\lambda_{g0}}{\lambda_{gc1} - \lambda_{gc2}} \right| \quad (33)$$

where λ_{g0} is the wavelength of resonance in the transmission line and λ_{gc} is the wavelength at the half power points. If the phase velocity in the transmission line does not vary with frequency, then the selectivity can be expressed simply in terms of either the wavelength or the frequency since

$$\frac{f}{f_0} - \frac{f_0}{f} = \frac{\lambda_0}{\lambda} - \frac{\lambda}{\lambda_0} \quad (34)$$

However, when the velocity in the transmission line varies with frequency, equation 34 does not hold true, and the expression relating the two parameters is more complicated. In the case of the rectangular waveguide

$$\lambda_g = \frac{c}{\sqrt{f^2 - f_{cw}^2}} \quad (35)$$

where c is the velocity of light in vacuum, f_{cw} is the cutoff frequency of the waveguide, $f_{cw} = \frac{c}{2a}$ and a is the width of the waveguide.

It can be shown readily that the frequency parameter can be expressed in terms of the wavelength, thus

$$\left(\frac{f}{f_0} - \frac{f_0}{f}\right) = \left(\frac{\lambda_{g0}}{\lambda_g} - \frac{\lambda_g}{\lambda_{g0}}\right) \left(\frac{\lambda_a}{\lambda_g}\right) \left(\frac{\lambda_{a0}}{\lambda_{g0}}\right) \quad (36)$$

where λ_a is the wavelength in free space.

For narrow percentage bands, this reduces to the approximate relation

$$\left(\frac{f}{f_0} - \frac{f_0}{f}\right) \doteq \left(\frac{\lambda_{g0}}{\lambda_g} - \frac{\lambda_g}{\lambda_{g0}}\right) \left(\frac{\lambda_{a0}}{\lambda_{g0}}\right)^2. \quad (37)$$

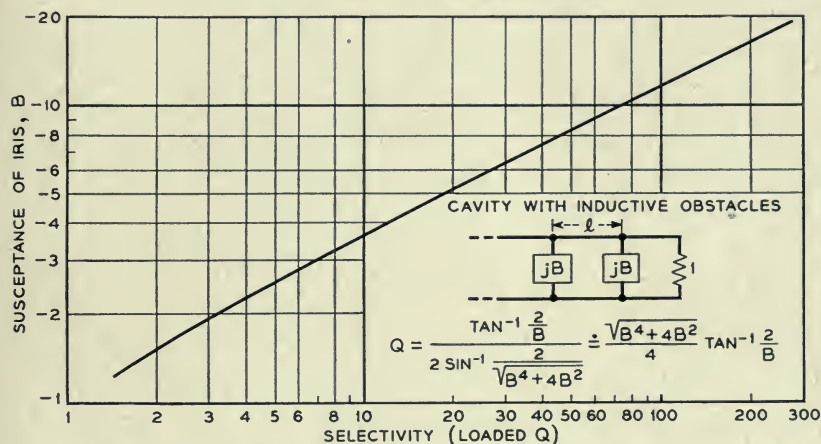


Fig. 11—The relation between loaded Q and normalized susceptance. (Inductive obstacles)

This states, in effect, that the percentage bandwidth is greater in terms of wavelength than in terms of frequency, by the square of the ratio of the wavelengths in the guide and in free space. The selectivity in terms of the frequencies and wavelength ratio thus becomes

$$Q \doteq \frac{1}{\left(\frac{f_z}{f_0} - \frac{f_0}{f_c}\right)} \cdot \left(\frac{\lambda_{a0}}{\lambda_{g0}}\right)^2 = \frac{f_0}{(f_{c2} - f_{c1})} \cdot \left(\frac{\lambda_{a0}}{\lambda_{g0}}\right)^2 \quad (38)$$

This is the selectivity that is plotted as a function of B in Figures 11 and 12.

EXCESS PHASE AND CONNECTING LINES

The excess phase of this type of cavity is taken into account by adding the lengths of line, ℓ' (see Fig. 9), which have a length given by the relation (See Appendix I)

$$\tan \frac{4\pi \ell'}{\lambda_{g0}} = - \left(\frac{j^2}{B}\right). \quad (39)$$

Combining Eq. 31 and Eq. 39 and solving for ℓ' in terms of ℓ

$$\ell' = \frac{\lambda_{g0}}{4} - \frac{\ell}{2} \quad (40)$$

where ℓ is the length of the cavity and λ_{g0} is the resonant wavelength in the line.

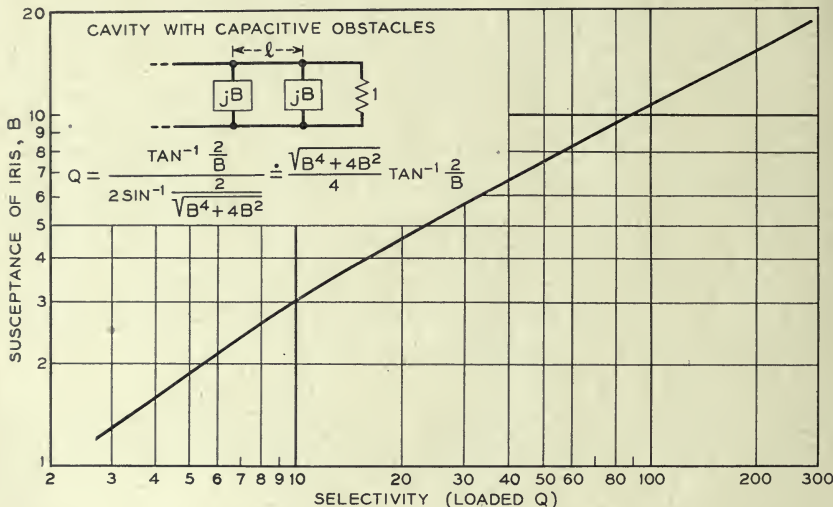


Fig. 12—The relation between loaded Q and normalized susceptance. (Capacitive obstacles)

Thus, when this length, corresponding to the excess phase of the cavity resonator, is absorbed in the length of line connecting two cavities together, the correct total connecting length becomes

$$\begin{aligned} \ell_s &= (2m + 1) \frac{\lambda_{g0}}{4} - \ell'_1 - \ell'_2 \\ &= \frac{\ell_1 + \ell_2}{2} - \frac{\lambda_g}{4} + m \frac{\lambda_g}{2} \end{aligned} \quad (41)$$

where ℓ_1 and ℓ_2 are the lengths of the cavities and m is any integer including zero.

OBSTACLES IN WAVEGUIDES

The three properties of the cavity—the resonant frequency, the selectivity and the excess phase—are given in Equations 31, 32 and 39, regardless of the sign of the normalized susceptance, B . In the case where the obstacles are inductive, B is negative; and where the obstacles are capacitive, B is positive.

Further explanation is needed to distinguish between these two important cases. First consider the case where inductive obstacles are used. $\tan \frac{2\pi\ell}{\lambda_{g0}}$ is negative and the cavity length lies between a quarter and a half wavelength (plus any multiple of half wavelength). The selectivity, as given by equation (32), is plotted on Fig. 11 for the fundamental mode. The excess phase is positive, and the added lengths, ℓ' , of Fig. 9 are positive. The connecting lines between two such cavities are then slightly less than a quarter wavelength (or odd multiple thereof).

Next consider the case where the obstacles are capacitive. $\tan \frac{2\pi\ell}{\lambda_{g0}}$ is positive and the cavity length lies between zero and a quarter wavelength (plus any multiple of half wavelengths). The selectivity as given by equation (32) is plotted in Fig. 12 for cavity lengths lying between a half wavelength and three quarters wavelength. The excess phase is negative and the added lengths, ℓ'_1 of Fig. 9 are negative. The connecting lines between two such cavities are then slightly longer than a quarter wavelength (or odd multiple thereof).

SUSCEPTANCE OF OBSTACLES

The Equations (31), (32) and (39) give the resonant wavelength, the selectivity (in terms of wavelength) and the excess phase as functions of the normalized susceptance of the obstacles which form the ends of the cavity, and a knowledge of this susceptance as a function of the geometrical configuration of the obstacle is necessary to complete the design of the filter. At low frequencies, conventional coils and condensers can be used to form the discontinuities in the transmission line; while at high frequencies, transmission line stubs can be used.¹⁴ In the microwave region, where waveguides are employed, obstacles having the shapes shown in Figures 13, 14, and 15 can be used.¹⁵

INDUCTIVE VANES

Figure 13 shows a plane metallic obstacle, transversely located across a rectangular waveguide, with a centrally located rectangular opening extending completely across the waveguide in a direction parallel to the electric vector. For thin obstacles, the normalized susceptance can be calculated from the approximate formula,¹⁵

$$B \doteq - \frac{\lambda_g}{a} \cot^2 \frac{\pi d}{2a} \quad (42)$$

where λ_g is the wavelength in the waveguide, a is the width of the waveguide, and d is the width of the iris opening.

When the iris is constructed of material of finite thickness, τ , the expression for the susceptance is more complicated,^{10, 11} and the equivalent circuit becomes a four-terminal network with both shunt and series elements. The equivalent shunt susceptance of this network can be obtained experimentally by measuring the insertion loss of the iris, from which a curve such as shown in Fig. 16 can be computed. These data* were taken for irises .050" thick in waveguide having internal dimensions of $0.872" \times 1.872"$ in the frequency range around 4000 mc. The ordinate is a parameter, K , from which the normalized susceptance is calculated:

$$B = K \left(\frac{\lambda_g}{2a} \right). \quad (43)$$

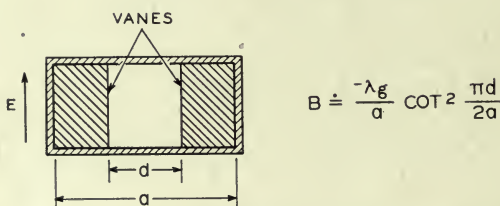


Fig. 13—One type of inductive obstacle in rectangular waveguide.

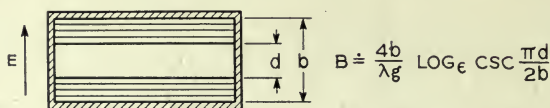


Fig. 14—One type of capacitive obstacle in rectangular waveguide.

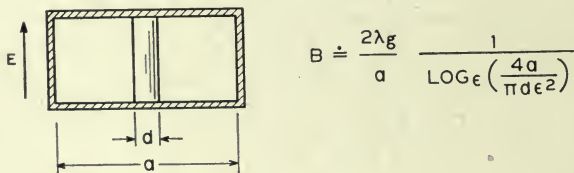


Fig. 15—Another type of inductive obstacle consists of a cylindrical post.

Along the abscissa is plotted the ratio of iris opening to width of the waveguide.

It can be demonstrated that for values of K from -1 to -20 , the equivalent iris opening is approximately the actual opening less the thickness of the metal sheet. For practical purposes, when the susceptance lies between -1.5 and -30 , it is often sufficient to use the approximation,

$$B \doteq -\frac{\lambda_g}{a} \cot^2 \left(\frac{\pi(d - \tau)}{2a} \right) \quad (44)$$

where τ is the thickness of the iris.

* Data supplied by Mr. L. C. Tillotson of Bell Telephone Laboratories.

CAPACITIVE IRISES

The normalized susceptance of infinitely thin capacitive obstacles, as illustrated in Fig. 14, may be calculated by the approximate relation¹⁵

$$B \doteq \frac{4b}{\lambda_g} \log_e \operatorname{cosec} \frac{\pi d}{2b} \quad (45)$$

where b is the height of the waveguide, λ_g is the wavelength in the waveguide, and d is the width of the iris opening.

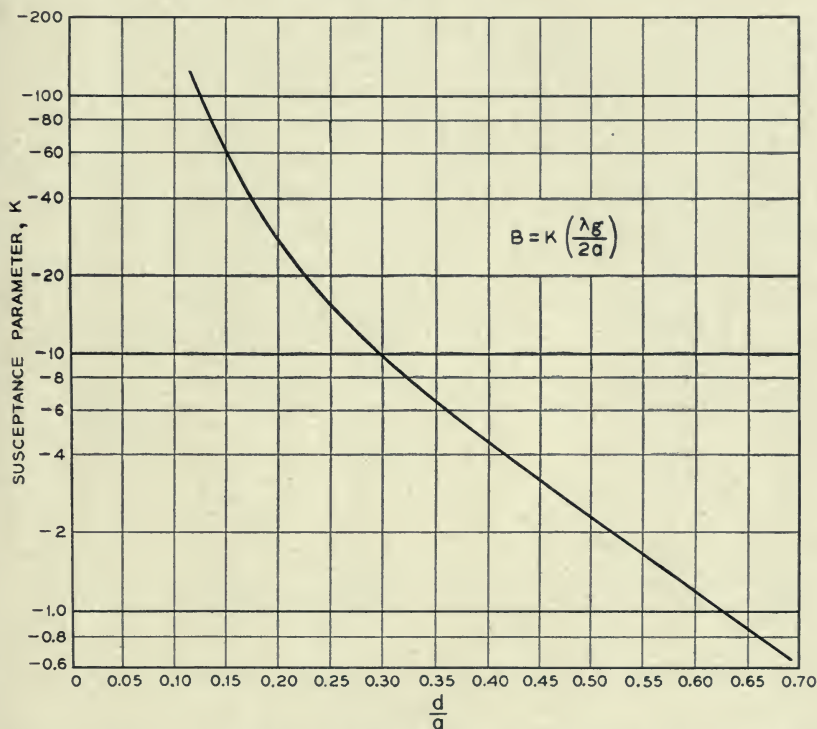


Fig. 16—Experimentally determined curve of normalized susceptance of inductive irises.

As with the inductive vanes, the normalized susceptance is a function of the iris thickness and may be calculated from the approximate formula¹⁵

$$B \doteq B_0 + \frac{2\pi\tau}{\lambda_g} \left(\frac{b}{d} - \frac{d}{b} \right) \quad (46)$$

where B_0 is the normalized susceptance of the infinitely thin iris, and τ is the iris thickness.

For best results, the irises should be designed from experimentally determined curves, however.

INDUCTIVE POSTS

The normalized susceptance of the round cylindrical inductive post, centrally located in the waveguide parallel to the electric vector, may be calculated from the approximate formula^{11, 15, 16}

$$B = -\frac{2\lambda_g}{a} \frac{1}{\log_e \left(\frac{4a}{\pi d \epsilon^2} \right)} \quad (47)$$

where a is the width of the guide, and d is the post diameter.

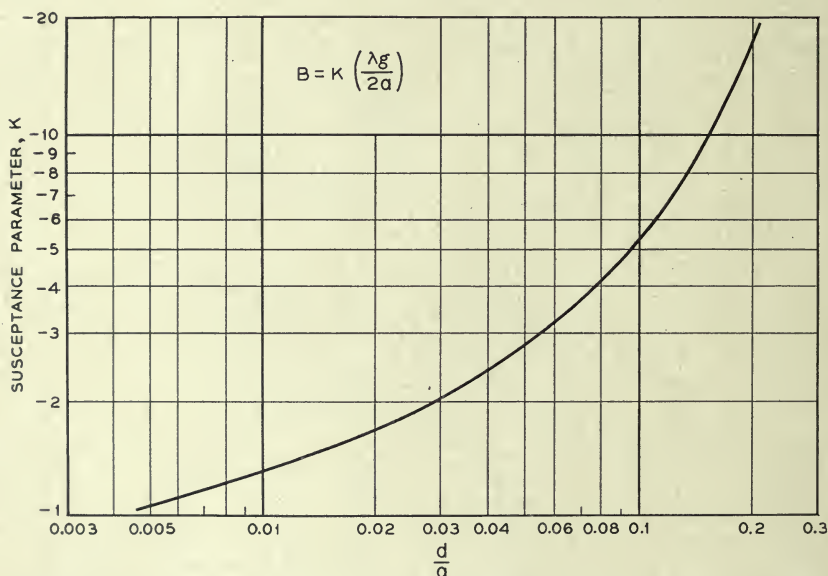


Fig. 17—Experimentally determined curve of normalized susceptance of inductive posts

The experimentally determined values of susceptance are somewhat less than the values calculated by the formula (47). The difference is less than 20% when $\frac{d}{a}$ is less than 0.08. A curve of experimentally determined values is plotted in Fig. 17, the data being taken in rectangular waveguide $0.872'' \times 1.872''$ at a frequency near 4000 mc.*

The normalized susceptance of posts is also a function of their position in the waveguide, the susceptance decreasing as the posts are moved off center. This feature may be used when it is desired to make all the posts in a filter

* Data supplied by Mr. A. E. Bowen of Bell Telephone Laboratories.

from stock of a given diameter. The expression for the normalized susceptance of off-center posts is given by the relation^{11, 16}

$$B \doteq -\frac{2\lambda_g}{a} \frac{1}{\sec^2 \frac{\pi s}{a} \left[\log_e \left(\frac{4a}{\pi d \epsilon^2} \cos \frac{\pi s}{a} \right) \right]} \quad (48)$$

where s is the distance off center.

EXPERIMENTAL DATA

The principles of waveguide filter design as outlined in the foregoing have been used in several applications. For example the channel branching filters in the New York-Boston microwave radio relay link consist of two resonant cavities separated by the equivalent of $\frac{3}{4}$ wavelength sections of waveguide. The transmitting modulators in this relay system also use two-chamber filters to separate the wanted sideband from the unwanted sideband. The transmission band in each of these applications was 10 mc

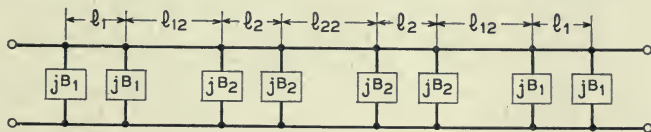


Fig. 18—Diagram of a transmission line filter consisting of four resonant cavities and three connecting lengths of line.

and the image frequency or the unwanted sideband which was to be reflected was 130 mc away.

In another case the requirements were that the standing wave ratio should be less than 0.64 db over a band of 20 mc and more than 28 db 30 mc on each side of the midband frequency. The design formulae indicated that a filter consisting of four cavities would be needed. These, then, would take the general configuration shown in Fig. 18, where the first and last cavities are formed by the obstacles jB_1 and the length of line ℓ_1 , while the two middle cavities are formed by the obstacles jB_2 and the length ℓ_2 . The lengths ℓ_{12} and ℓ_{22} correspond to the transforming sections of transmission line which connect the cavities together. The loaded Q 's required to meet the specifications turned out to be $Q_1 = 12.25$ and $Q_2 = 30.0$, after allowance had been made for the selectivities of the $\frac{3}{4}$ wavelength connecting sections. Assuming that the cavities would be formed with inductive obstacles, as shown schematically in Fig. 19, the susceptances to obtain these selectivities were obtained from Fig. 11 based on equation 32. This gave

$$B_1 = -4.08$$

$$B_2 = -6.36$$

These susceptances were realized with centrally located round posts, for which the data of Fig. 17 has been plotted, and this filter was constructed according to the calculated dimensions which are shown in Fig. 20. Each of the four cavities was tuned separately to resonance near midband by adjusting a capacitive plug located in the center of each. The characteristic then obtained is plotted in Fig. 21, which shows that the standing wave

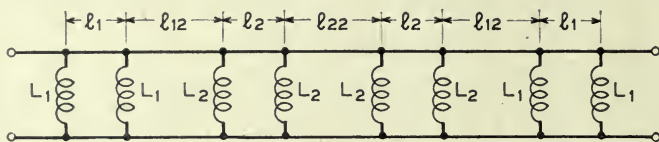


Fig. 19—A four-cavity filter which utilizes inductive obstacles.

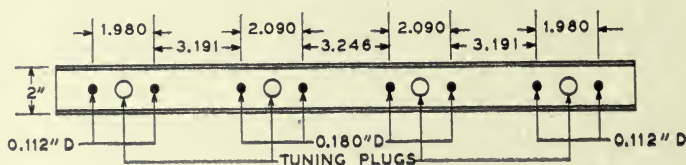


Fig. 20—The calculated dimensions for a four-cavity maximally-flat filter in 0.872" x 1.872" rectangular waveguide.

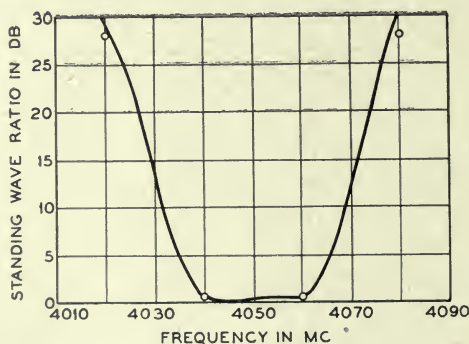


Fig. 21—Measured characteristic of four-cavity filter of Figure 20.

ratio met the design points quite well. These are shown as circles in the figure. The insertion loss of this filter was less than 0.7 db over a 25-mc band and less than 0.3-db at midband.

Another maximally-flat waveguide filter consisting of fifteen resonant cavities gave an insertion loss of two decibels at midband, 4-db loss at 20-mc bandwidth and 40-db loss at 30-mc bandwidth. The input standing wave ratio was less than 1.0 db over a 20-mc band. Its characteristics are plotted in Figs. 22 and 23. This excellent performance is remarkable in

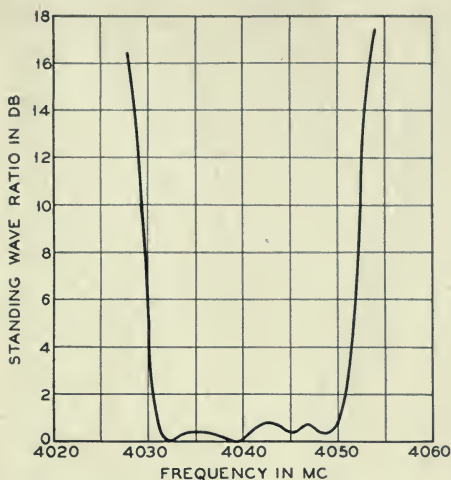


Fig. 22—Measured standing wave ratio of maximally-flat filter consisting of fifteen resonant cavities.

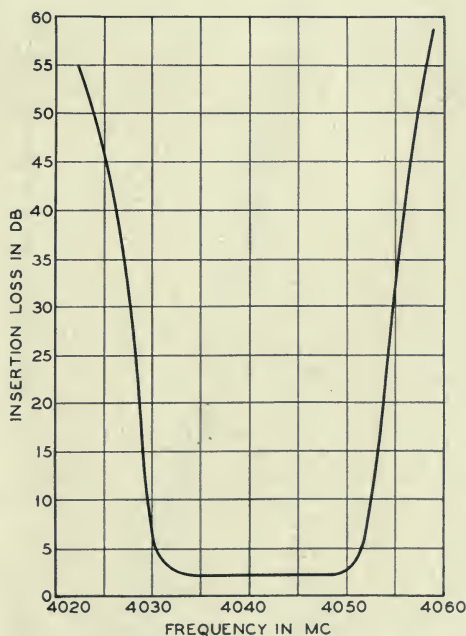


Fig. 23—Measured insertion loss of the fifteen-cavity filter.

view of the difficulties that might be encountered in constructing and aligning a filter consisting of 75 discontinuities and 29 lengths of waveguide. Its physical length (over 80") may be seen in the photograph of Fig. 24.

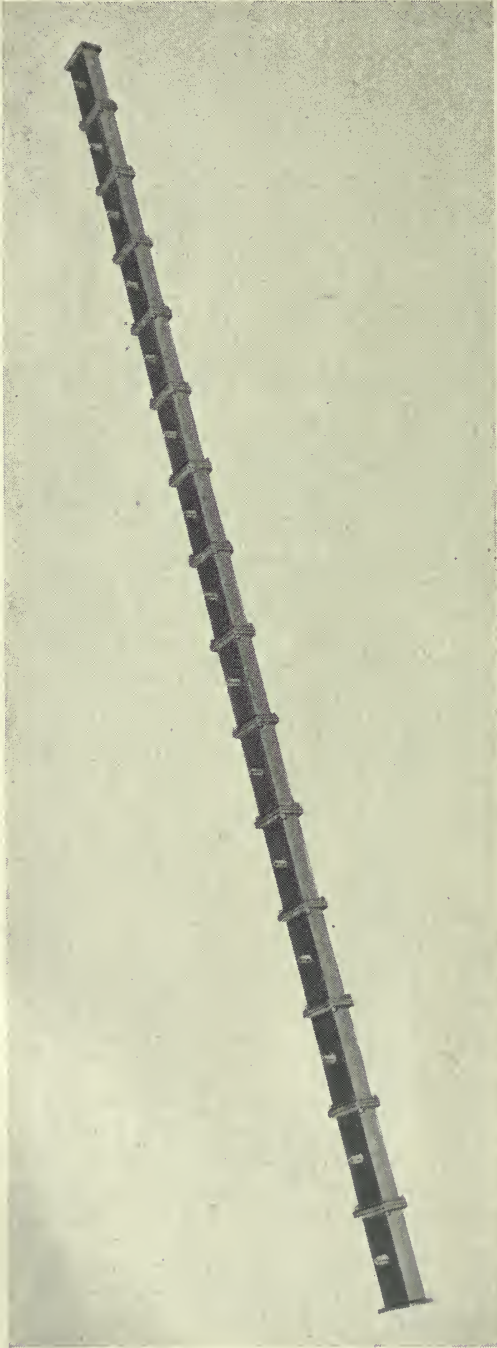


Fig. 24—The fifteen cavity filter.

The theoretical treatment of maximally-flat filters presented here has ignored the dissipation in the elements. Better agreement between expected and observed characteristics would be obtained if this had been taken into account. The observation of .3-db loss and 2-db loss in the four-cavity and the fifteen-cavity maximally-flat filters is indicative of the amounts of added insertion loss to be expected because of dissipation in the elements. In addition to the increased loss at midband, we should expect a rounding of the insertion loss characteristic near the cutoff frequencies, and a broadening of the standing wave characteristic at frequencies well beyond cutoff. In many applications, however, these effects can be ignored.

CONCLUDING REMARKS

In the foregoing, the design of maximally-flat band-pass filters has been treated in detail. The treatment of other types of band-pass and band-rejection filters is beyond the scope of the present paper, although much of the material presented here may be of use in designing such filters. In fact, almost any filter consisting of a ladder network of inductive and capacitive elements in series and in shunt can be simulated in waveguides by following these principles. Emphasis on the maximally-flat filter has been deliberate for two reasons: it gives a type of transmission characteristic that is useful in microwave work; it is simple to design.

ACKNOWLEDGEMENT

Many members of the Holmdel Radio Research Laboratories have influenced the evolution of the technique of building waveguide filters. A firm foundation was laid by the pioneering work of G. C. Southworth¹⁷ and A.G. Fox. Intimate association with W. D. Lewis and L. C. Tillotson fostered many stimulating discussions which clarified doubtful issues. The comments of M. D. Brill and S. Darlington are deeply appreciated. The untiring effort of R. H. Brandt in making skillfully the accurate measurements which were necessary is gratefully acknowledged as is also the invaluable support of all the many other people without whose cooperation and confidence the work would have been impossible.

REFERENCES

1. E. L. Norton, Constant Resistance Networks with Applications to Filter Groups, *B. S. T. J.*, Vol. XVI, pp. 178-193, April 1937.
2. W. D. Lewis and L. C. Tillotson, *B. S. T. J.*, Vol. XXVII, #1, pp. 83-95, January 1948.
3. E. L. Norton, U. S. Patent #1,788,538, January 13, 1931.
4. W. R. Bennett, U. S. Patent #1,849,656, March 15, 1932.
5. S. Butterworth, *Wireless Engineer*, Vol. VII, #85, pp. 536-541, October 1930.
6. V. D. Landon, *R.C.A. Review*, Vol. V, pp. 347-362 and pp. 481-497, 1940-41.
7. Henry Wallman, *M.I.T. Radiation Laboratory Report* #524, February 23, 1944.
8. S. Darlington, *Journal of Math. and Phys.*, Vol. XVIII, pp. 257-353, September 1939.

9. R. M. Fano and A. W. Lawson, *Proc. I.R.E.*, Vol. 35, #1, pp. 1318-1323, November 1947.
10. Wilbur L. Pritchard, *Journal of Applied Physics*, Vol. XVIII, #10, pp. 862-872, October 1947.
11. M.I.T. Wave Guide Handbook, *Supplement Rad. Lab. Rpt. 41*, Sec. 21a, (January 23, 1945).
12. A. L. Samuel, J. W. Clark and W. W. Mumford, Gas Discharge Transmit-Receive Switch. *B. S. T. J.*, Vol. XXV, #1, pp. 48-101, January 1946.
13. S. A. Schelkunoff, Representation of Impedance Functions in Terms of Resonant Frequencies. *Proc. I.R.E.*, Vol. XXXII, pp. 83-90, February 1944.
14. W. P. Mason and R. A. Sykes, *B. S. T. J.*, Vol. XVI, pp. 275-302, July 1937.
15. Sperry Microwave Transmission Design Data, Publication No. 23-80, Sperry Gyroscope Co.
16. S. A. Schelkunoff, *Quarterly of Applied Math.*, Vol. I, #1, pp. 78-85, April 1943.
17. G. C. Southworth, Hyper-Frequency Wave Guides, *B. S. T. J.*, Vol. XV, pp. 284-309, April 1936.

APPENDIX I

A cavity resonator, consisting of a length of transmission line, ℓ , at each end of which there is an unvarying susceptance, jB , is approximately equivalent to a tuned circuit, consisting of an inductance, L , and a capacity, C , in parallel located at the center of a short length of transmission line, $2\ell'$, when these two conditions are satisfied:

(1) The square root of L over C is equal to the surge impedance of the transmission line divided by twice the loaded Q of the cavity.

(2) The sum of the lengths of the two transmission lines ℓ and $2\ell'$ is equal to a half wavelength at resonance.

The first of these conditions follows from equation 3 of the text above, and the proof of the second condition will be given in the following analysis, based on the schematic drawing of Figures 9 and 10. In this analysis, the loaded Q of the cavity is derived in terms of the susceptance of the obstacles at its ends.

Since the cavity and the tuned circuit are both symmetrical it is adequate to consider but one half of each in establishing the equivalence. Then by setting the short circuit admittance of one equal to the other and setting the open circuit admittance of one equal to the other, the necessary relationships are derived.

The following symbols will be used in addition to those used in the text:

Y_{sc} = Normalized admittance, short circuited.

Y_{oc} = Normalized admittance, open circuited.

The subscripts 1 and x refer to the cavity and the equivalent tuned circuit respectively.

$$\theta_1 = \frac{2\pi}{\lambda_g} \cdot \frac{\ell}{2}$$

$$\theta_x = \frac{2\pi}{\lambda_g} \cdot \ell'$$

The short-circuited admittances of half the cavity and half the tuned circuit are

$$Y_{sc1} = j(B_1 - \cot \theta_1) \quad (\text{A1})$$

$$Y_{scx} = -j \cot \theta_x \quad (\text{A2})$$

while the open-circuited admittances are

$$Y_{oc1} = j(B_1 + \tan \theta_1) \quad (\text{A3})$$

$$Y_{ocx} = \frac{j\left(\frac{B_x}{2} + \tan \theta_x\right)}{1 - \frac{B_x}{2} \tan \theta_x} \quad (\text{A4})$$

Putting $Y_{sc1} = Y_{sc2}$

$$\tan \theta_x = \frac{1}{\cot \theta_1 - B_1} \quad (\text{A5})$$

Putting A5 in A4 and setting $Y_{oc1} = Y_{ocx}$

$$B_1 + \tan \theta_1 = \frac{\frac{B_x}{2} + \frac{1}{\cot \theta_1 - B_1}}{1 - \frac{B_x}{2} \frac{1}{\cot \theta_1 - B_1}} \quad (\text{A6})$$

Solving for B_x we have

$$B_x = -B_1(B_1 \sin 2\theta_1 - 2 \cos 2\theta_1) \quad (\text{A7})$$

which becomes

$$B_x = \sqrt{B_1^4 + 4B_1^2} \sin \frac{2\pi\ell}{\lambda_{g0}} \left(\frac{\lambda_{g0}}{\lambda_g} - 1 \right) \quad (\text{A8})$$

where

$$\frac{2\pi\ell}{\lambda_{g0}} = \arctan \frac{2}{B_1} \quad (\text{A9})$$

Equation A9 gives the requirements for resonance.

The expression for the loaded Q is

$$Q = \frac{\frac{2\pi\ell}{\lambda_{g0}}}{\frac{2\pi\ell}{\lambda_{gc2}} - \frac{2\pi\ell}{\lambda_{gc1}}} \quad (\text{A10})$$

The cutoff wavelengths are obtained when B_x is equal to ± 2 and we have from equation A8

$$\frac{2\pi\ell}{\lambda_{gc1}} = \frac{2\pi\ell}{\lambda_{g0}} - \arcsin \frac{2}{B_1\sqrt{B_1^2 + 4}}, \quad (\text{A11})$$

$$\frac{2\pi\ell}{\lambda_{gc2}} = \frac{2\pi\ell}{\lambda_{g0}} + \arcsin \frac{2}{B_1\sqrt{B_1^2 + 4}}, \quad (\text{A12})$$

from which we obtain

$$Q = \frac{\arctan \frac{2}{B}}{2 \arcsin \frac{2}{\sqrt{B^4 + 4B^2}}} \doteq \frac{\sqrt{B_1^4 + 4B_1^2}}{4} \arctan \frac{2}{B}. \quad (\text{A13})$$

This gives the loaded Q of the cavity in terms of the susceptance of the end obstacles.

To derive the length corresponding to the excess phase of the cavity, let the short-circuited admittances be equal by equating equations A1 and A2, and let the wavelength be the resonant wavelength of the cavity, and we have

$$B_1 - \cot \theta_{10} = -\cot \theta_{x0}. \quad (\text{A14})$$

From equation A9

$$B_1 = 2 \cot 2\theta_{10} \quad (\text{A15})$$

so that

$$2 \cot 2\theta_{10} - \cot \theta_{10} = -\cot \theta_{x0}. \quad (\text{A16})$$

But

$$2 \cot 2\theta_{10} - \cot \theta_{10} = -\tan \theta_{10} \quad (\text{A17})$$

hence

$$\tan \theta_{10} = \cot \theta_{x0} \quad (\text{A18})$$

or

$$\theta_{10} + \theta_{x0} = \frac{\pi}{2} \quad (\text{A19})$$

That is

$$\frac{2\pi}{\lambda_{g0}} \cdot \frac{\ell}{2} + \frac{2\pi\ell'}{\lambda_{g0}} = \frac{\pi}{2}$$

whence

$$\ell + 2\ell' = \frac{\lambda_{g0}}{2} \quad (\text{A20})$$

which proves the second condition mentioned above, namely, that the sum of the lengths of the transmission lines in the cavity and its equivalent circuit is equal to a half wavelength.

The normalized admittance of the cavity terminated in the surge admittance of the guide can be written in terms of its loaded Q and a wavelength variable as

$$YR \doteq 1 + j2Q \left[2 \left(\frac{\lambda_{g0}}{\lambda_g} \right) - 1 \right]. \quad (\text{A21})$$

This expression is obtained from equations A8 and A13 by making the assumption that the bandwidth is narrow so that the sine of the angle in equation A8 can be replaced by the angle. This admittance is referred to a point slightly inside the cavity, i.e. a distance ℓ' inside.

The similarity between this expression and the corresponding one for the parallel resonant circuit consisting of lumped elements is evident. (See eq. 7 of the text.)

$$YR = 1 + j2Q \left[\frac{f}{f_0} - \frac{f_0}{f} \right] \quad (\text{A22})$$

In the case of the cavity the bracketed term is a wavelength variable; in the case of the tuned circuit it is a frequency variable.

The loss function for maximally-flat filters in waveguides becomes

$$\frac{P_0}{P_L} \doteq 1 + \left[Q_T 2 \left(\frac{\lambda_{g0}}{\lambda_g} - 1 \right) \right]^{2n}. \quad (\text{A23})$$

The loaded Q 's of the cavities taper sinusoidally from one end of the filter to the other so that

$$Q_r = Q_T \sin \left(\frac{2r-1}{2n} \right) \pi. \quad (\text{A24})$$

Transient Response of an FM Receiver

By MANVEL K. ZINN

INTRODUCTION

THIS paper develops various formulas for the response of an FM receiver to signal or noise input voltages of arbitrary form. The principal object in view is to obtain a more complete understanding of how an FM receiver responds to transient voltages, such as those arising from ignition interference, but the more general aspects of the theory have other applications as well. In particular, general formulas are given for the response of a linear circuit to an applied voltage, or current, of variable frequency. The Fourier transforms, or frequency spectra, of the response, and the envelope thereof, are determined.

Two examples are given: (1) the audio response of an FM receiver to a very large impulse and (2) the response, including harmonic distortion, to a sinusoidal signal wave.

The element of an FM receiver that demands most discussion is the balanced frequency detector. The greater part of the paper accordingly deals with that important element. The general problem can be stated as follows: A limiter and frequency detector are transmitting a steady unmodulated carrier wave to an audio output circuit. At time, $t = 0$, frequency modulation of arbitrary form is applied to the carrier (either by signal modulation or a superposed noise transient). What is the audio output voltage that results?

FREQUENCY DETECTOR

Except for the greater bandwidth, the amplifiers and selective circuits between the antenna and the limiter of an FM receiver are similar to those of an AM receiver in their transmission features. If the selective circuits have a bandwidth ample to accommodate the maximum frequency swing of the FM transmitter, and if the transmission over the band is substantially "flat" and the phase shift nearly linear with frequency, the amplifiers will introduce little distortion. The limiter and frequency detector are therefore regarded as the distinctive elements of an FM receiver meriting theoretical discussion.

The literature contains descriptions of frequency detectors of several types together with adequate analyses of the action of the circuits based on the variable impedance concept.¹ The more generally used circuits can be

¹ See Items 3 to 6 in list of references attached.

reduced to the circuit shown schematically in Fig. 1, which can be taken to illustrate a generic form of frequency detector. Z_1 and Z_2 are two resonant impedances tuned to different frequencies, one above, the other below, the carrier frequency.² For example, the simplest version of Z_1 and Z_2 could be, for each, a parallel combination of R , L and C . Across each of these impedances is connected a rectifier with load circuit so proportioned that the rectification is substantially linear. The rectifiers are poled so that their low-frequency outputs are opposed, thereby obtaining cancellation of even-order demodulation products. With this arrangement, the low-frequency output voltage V_o , which is applied to the audio amplifier, is substantially proportional to the difference between the envelopes of the voltage drops across Z_1 and Z_2 .

The resistance elements of the impedances, Z_1 and Z_2 , each include a shunting resistance equal to half the load resistance of the associated rectifier, which therefore determines, to some extent, the Q of the tuned circuit. The output diode load, $R_o C_o$, has negligible impedance at the carrier frequency. Under these conditions, the low-frequency output voltage across the two-rectifier load impedances is

$$V_o = \eta ([V_1] - [V_2])$$

where η = detection efficiency (nearly unity)

V_1, V_2 = high-frequency voltages across Z_1, Z_2 , respectively (Fig. 1)

$[V]$ = envelope of V .

All this is in accord with the accepted understanding of the operation of a properly designed linear rectifier working at an efficiency approaching 100 per cent.

The amplitudes of the voltages, V_1 and V_2 , across the resonant impedances, Z_1 and Z_2 , of Fig. 1 are shown in Fig. 2. In the practical engineering analysis of this frequency detector circuit, employing the idea of impedance that varies in step with the instantaneous frequency, the two voltages of Fig. 2 are subtracted (owing to the opposed polarities of the rectifiers) to obtain the over-all voltage-frequency characteristic shown in Fig. 3. Then it is inferred, by physical intuition, that if the instantaneous frequency of the carrier is varied at the input, the output voltage wave will vary as indicated by the curve of Fig. 3. Strictly speaking, this is a false assumption, but where the rate of variation of the instantaneous frequency is at an audio signal frequency far below the carrier frequency, the error in the assumption is of no importance, whereas the simplification in thinking accomplished

² The term *carrier frequency* will be used to designate the value of the unmodulated received frequency after all heterodyne conversions. (This frequency is equal to the mid-band frequency of the last intermediate frequency amplifier ahead of the limiter, if tuning is perfect.)

by it is considerable. It is only where the rate of variation of the instantaneous frequency is high, as it can be in the case of a large noise transient caused by impulse excitation, that the error in the assumption in question

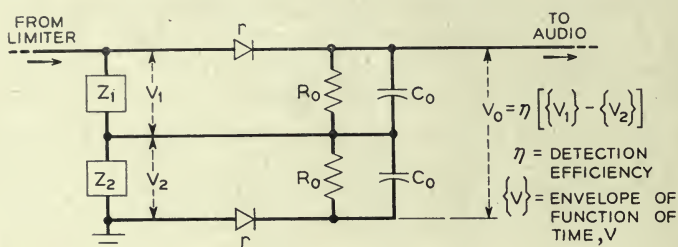


Fig. 1—Circuit of a balanced frequency detector.

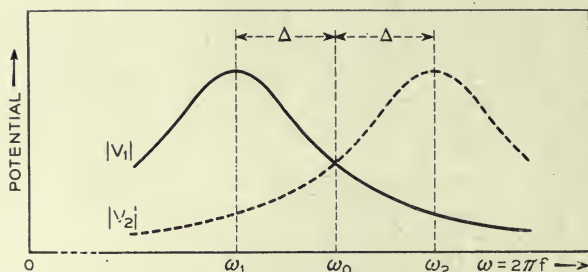


Fig. 2—Voltages across tuned circuits of frequency detector.

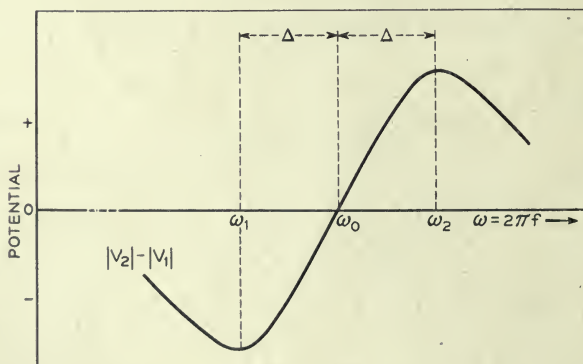


Fig. 3—Output voltage of frequency detector.

can become serious. A particularly subtle error that can arise from the assumption is to fall into the habit of regarding the characteristic curve of Fig. 3 as a frequency transmission curve of the sort obtained by measuring the ratio of output to input of a linear network over a range of frequencies.

The curve of Fig. 3 is not such a transmission curve, because the principle of superposition does not apply and a frequency conversion is involved.

Owing to the considerations discussed above, the analysis to follow avoids the assumption of variable impedance associated with the varying instantaneous frequency. This does not imply that the assumption, as employed by various writers, is considered seriously erroneous, but, rather, that it seems preferable to develop the theory without invoking the assumption, provided that this can be done without falling into unmanageable complications. Briefly, the procedure in the work to follow is to determine directly the envelopes of the voltages V_1 and V_2 as functions of time, one envelope then being subtracted from the other to obtain the output wave.

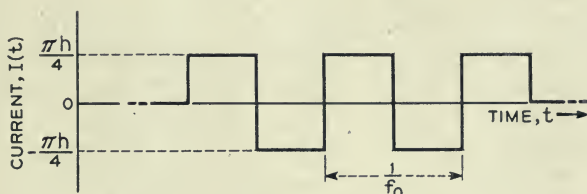


Fig. 4—Current wave out of limiter.

GENERAL THEORY

When a carrier is being received, the limiter can be regarded as substantially a constant current source having an internal shunt admittance small compared to the admittance of the tuned impedance elements, Z_1 and Z_2 , of the frequency detector. If the limiting is severe, as it should be for good operation, the current delivered by the limiter is a rectangular wave as illustrated in Fig. 4. When this current is driven through the impedances, Z_1 and Z_2 , the voltage drops, V_1 and V_2 , that arise across these elements are substantially sinusoidal in form, owing to the selectivity, which practically extinguishes all harmonics of the carrier frequency. We therefore take the current input to be sinusoidal in the first place, namely

$$I(t) = h \cos 2\pi f_0 t$$

This is the unmodulated current, $\pi h/4$ being the current cutoff point of the limiter and f_0 the frequency of the carrier. When the carrier is modulated in frequency, we write

$$I(t) = h \cos [2\pi f_0 t + \theta(t)] \quad (1)$$

where $\theta(t)$ is the phase angle varying with time. The instantaneous frequency then is

$$f(t) = \frac{1}{2\pi} \frac{d}{dt} [2\pi f_0 t + \theta(t)] = f_0 + \frac{1}{2\pi} \theta'(t). \quad (2)$$

In the transmission of signals by frequency modulation, the instantaneous radian frequency deviation, $\theta'(t)$, is made to vary in proportion to the signal amplitude, so that $\theta(t)$ then varies in proportion to the time integral of the signal amplitude.

As a preliminary step to the discussion of the frequency detector itself, we require a formula for the voltage drop across an impedance $Z(f)$ when the frequency-modulated current (1) flows through it. The point of view usually adopted is to regard the impedance as a composite function of time, viz., $Z[f(t)]$, and to say that the voltage across it is

$$V(t) = I(t)Z[f(t)] = I(t)Z\left[f_0 + \frac{1}{2\pi} \theta'(t)\right]. \quad (3)$$

This quasi-stationary viewpoint gives results that are nearly correct if the rate of change, $\theta''(t)/2\pi$, of the variable frequency is not too large. The magnitude of the error has been determined in a paper by Carson and Fry.³ In the present paper, impedance is a function of frequency that is independent of time, as in the classic theory of linear systems. The frequent use of the term "instantaneous frequency," as defined by (2), does not imply a departure from this point of view.

In the following, $H(t)$ is, in general, the voltage response as a function of time, of a network to a unit impulse of current applied at time $t = 0$. In the case of a two-terminal impedance element, $H(t)$ is the voltage drop across the element when a unit impulse of current is sent through it. Then, if the frequency modulated current (1) flow through the impedance, the voltage drop is

$$V(t) = h \int_0^\infty \cos [\omega_0(t - \tau) + \theta(t - \tau)] H(\tau) d\tau \quad (4)$$

where $\omega_0 = 2\pi f_0$. $\theta(t)$ can have any form as a function of time. $V(t)$ can be written,

$$\begin{aligned} V(t) &= \frac{h}{2} e^{i\omega_0 t} \int_0^\infty e^{-i\omega_0 \tau + i\theta(t-\tau)} H(\tau) d\tau \\ &\quad + \frac{h}{2} e^{-i\omega_0 t} \int_0^\infty e^{i\omega_0 \tau - i\theta(t-\tau)} H(\tau) d\tau. \end{aligned} \quad (5)$$

In the frequency detector problem, the result finally desired is the envelope of the voltage wave. It will clarify the discussion to explain first what is meant by an envelope. If the voltage is of the form

$$\begin{aligned} V(t) &= c(t) \cos [\omega_0 t + \phi(t)] \\ &= a(t) \cos \omega_0 t - b(t) \sin \omega_0 t \\ &= \frac{1}{2}[a(t) + ib(t)] e^{i\omega_0 t} + \frac{1}{2}[a(t) - ib(t)] e^{-i\omega_0 t} \end{aligned} \quad (6)$$

³ Item 1 in the bibliography. See formula 21 in that paper.

the complex function,

$$[V(t)] = c(t) e^{i\phi(t)} = a(t) + ib(t) \quad (7)$$

is here called the "envelope function" of the voltage with respect to the radian frequency ω_0 , $c(t)$ being a real amplitude modulation factor, which is the envelope⁴ itself, as usually conceived, and $\exp [i\phi(t)]$ a complex frequency modulation factor, in which $\phi'(t)$ is the instantaneous deviation of the radian frequency from the reference value, ω_0 . If such a modulated voltage wave is applied to an ideal linear detector, the output voltage across the load circuit of the latter is the real envelope, $c(t) = [a^2(t) + b^2(t)]^{1/2}$. This concept of an envelope function provides a convenient generalization of modulation ideas. Both amplitude modulation and frequency modulation vary the envelope function, but in different ways. In amplitude modulation, the real magnitude, $c(t)$, is varied while the angle ϕ is constant, whereas, in frequency modulation, c is constant and it is the angle, $\phi(t)$, that is varied.

It will be seen that (5) is in precisely the same form as (6), so that we can write the envelope function of $V(t)$ immediately, as follows:

$$[V(t)] = a(t) + ib(t) = h \int_0^\infty e^{-i\omega_0\tau + i\theta(t-\tau)} H(\tau) d\tau. \quad (8)$$

The conjugate envelope function then is

$$[\overline{V(t)}] = a(t) - ib(t) = h \int_0^\infty e^{i\omega_0\tau - i\theta(t-\tau)} H(\tau) d\tau. \quad (9)$$

The spectrum of the envelope function is also of interest. To obtain the spectrum, which we shall call, $F_0(f)$, we find the Fourier transform (hereafter abbreviated, F.T.) of both sides of (8), viz.:

$$F_0(f) = \int_{-\infty}^\infty [V(t)] e^{-i\omega t} dt = h \int_{-\infty}^\infty e^{-i\omega t} \int_0^\infty e^{-i\omega_0\tau + i\theta(t-\tau)} H(\tau) d\tau dt. \quad (10)$$

It is permissible to reverse the order of integration of τ and t , obtaining

$$F_0(f) = h \int_0^\infty e^{-i\omega_0\tau} H(\tau) \int_{-\infty}^\infty e^{-i\omega t + i\theta(t-\tau)} dt d\tau. \quad (11)$$

The F.T. of $h \exp [i\theta(t)]$ will be designated, $\Psi(f)$, i.e.

$$\Psi(f) = h \int_{-\infty}^\infty e^{i\theta(t) - i\omega t} dt. \quad (12)$$

⁴ The "envelope", so defined, is an engineering concept and is not quite the same thing as the envelope of mathematics, which is always tangent to a curve or set of curves.

Putting $t - \tau$ in place of t in place of t as the variable of integration in (12) we have

$$\Psi(f) = h e^{i\omega\tau} \int_{-\infty}^{\infty} e^{-i\omega t + i\theta(t-\tau)} dt. \quad (13)$$

Thus it is seen that the inner integral of (11) is equal to $e^{-i\omega\tau}\Psi(f)$ and the equation becomes

$$F_0(f) = \Psi(f) \int_0^{\infty} H(\tau) e^{-i(\omega+\omega_0)\tau} d\tau. \quad (14)$$

Now the F.T. of $H(t)$ is $Z(f)$, i.e.

$$Z(f) = \int_{-\infty}^{\infty} H(t) e^{-i\omega t} dt. \quad (15)$$

Therefore

$$Z(f + f_0) = \int_{-\infty}^{\infty} H(t) e^{-i(\omega+\omega_0)t} dt. \quad (16)$$

This differs from the integral in (14) only in the lower limit of integration. But since $H(t)$ is the response to an impulse applied at time $t = 0$, $H(t) = 0$ for $t < 0$ and the two integrals are therefore equal. Putting (16) in (14) we have, finally

$$F_0(f) = \int_{-\infty}^{\infty} [a(t) + ib(t)] e^{-i\omega t} dt = \Psi(f) Z(f + f_0). \quad (17)$$

The F.T. of the conjugate envelope function, $a - ib$, is

$$\bar{F}_0(-f) = \int_{-\infty}^{\infty} [a(t) - ib(t)] e^{-i\omega t} dt = \bar{\Psi}(-f) \bar{Z}(-f + f_0) \quad (18)$$

where symbols with the superbar denote the complex conjugates of unbarred symbols. Since $Z(f)$ is the F.T. of a real variable, $H(t)$, it must assume conjugate values for positive and negative values of f , i.e., $Z(f) = \bar{Z}(-f)$ and therefore $Z(f + f_0) = \bar{Z}(-f + f_0)$. Consequently, (18) could be written

$$\bar{F}_0(-f) = \bar{\Psi}(-f) Z(f + f_0) \quad (19)$$

(17) and (18) are the final solutions in frequency functions corresponding to the solutions (8) and (9) in time functions. The formulas in frequency functions have the advantage of compactness, which makes them easy to remember.

We require also the F.T. of the voltage itself, $v(t)$, which we shall call $F(f)$. From (6)

$$F(f) = \int_{-\infty}^{\infty} V(t) e^{-i\omega t} dt = \frac{1}{2} \int_{-\infty}^{\infty} (a + ib) e^{-i(\omega-\omega_0)t} dt \\ + \frac{1}{2} \int_{-\infty}^{\infty} (a - ib) e^{-i(\omega+\omega_0)t} dt \quad (20)$$

and from (17) and (18) this evidently is

$$F(f) = \frac{1}{2}\Psi(f - f_0)Z(f) + \frac{1}{2}\bar{\Psi}(-f - f_0)\bar{Z}(-f) \quad (21)$$

or, since $Z(f) = \bar{Z}(-f)$

$$F(f) = \frac{1}{2}Z(f) [\Psi(f - f_0) + \bar{\Psi}(-f - f_0)]. \quad (22)$$

Anyone familiar with the rules of Fourier transforms could write down this frequency function in the first place and then proceed to find the time functions by the reverse of the process just carried out. But the time functions are more closely related to the physics of the problem and therefore provide a more fundamental starting point for its solution.

It will be appreciated that, although the above discussion has been phrased to apply to the problem of finding the voltage drop across an impedance when a frequency modulated current flows through it, the formulas also give the current through an admittance when a frequency-modulated voltage is applied across it. They also give the output voltage or current of a four-terminal network when a frequency-modulated current or voltage is applied at the input. These various applications of the formulas obviously can be made by placing definitions on Z and H appropriate to the particular problem.

The next step is to assume suitable values of the impedance $Z(f)$ and various forms of the frequency modulation function θ and to employ these particular values in the general formulas 8, 9, 17 and 18.

BALANCED FREQUENCY DETECTOR

The impedance can be defined, in general, as a rational algebraic function, viz.:

$$Z(i\omega) = \frac{(i\omega - a_1)(i\omega - a_2) \cdots (i\omega - a_m)}{(i\omega - p_1)(i\omega - p_2) \cdots (i\omega - p_n)} = \frac{P(i\omega)}{Q(i\omega)}. \quad (23)$$

Writing the polynomials P and Q in this way as the products of their factors exhibits the a 's as the zeros of Z and the p 's as the poles. The latter determine the frequencies of free vibration of the network. For the network to be stable, the p 's must all have negative real parts. The a 's and p 's are either real or occur in conjugate complex pairs. By the partial fraction rule, the expression can be broken up into a series of simple fractions; thus

$$Z(i\omega) = \frac{A_1}{i\omega - p_1} + \frac{A_2}{i\omega - p_2} + \cdots. \quad (24)$$

If the poles are all simple, the A 's are given by

$$A_j = \frac{P(p_j)}{Q'(p_j)}. \quad (25)$$

For the present purpose, only a pair of terms of (24) need be considered. This will provide a specific solution of the balanced frequency detector for the case, previously used for illustration, where the impedances are two simple resonant circuits of parallel R , L and C . At the same time, this solution can be extended to more complicated circuits by superposing a number of such elementary solutions, as is clearly possible with the type of impedance development indicated by (24).

The impedance of R , L and C in parallel, written in the form of (23) and (24), is

$$Z(i\omega) = \frac{i\omega}{C} \frac{1}{(i\omega - p_1)(i\omega - p_2)} = \frac{k}{i\omega - p_1} + \frac{\bar{k}}{i\omega - p_2} \quad (26)$$

$$= \frac{k}{i\omega + \gamma} + \frac{\bar{k}}{i\omega + \bar{\gamma}}$$

where

$$-p_1 = \gamma = \alpha + i\beta$$

$$-p_2 = \bar{\gamma} = \alpha - i\beta$$

$$k = \frac{1}{2C} \left(1 - \frac{i\alpha}{\beta} \right)$$

$$\bar{k} = \frac{1}{2C} \left(1 + \frac{i\alpha}{\beta} \right)$$

$$\alpha = \frac{1}{2RC}, \quad \beta = \sqrt{\frac{1}{LC} - \frac{1}{4R^2C^2}} = \sqrt{\omega_c^2 - \alpha^2}, \quad \omega_c = \frac{1}{\sqrt{LC}}.$$

The voltage response of the circuit to a unit impulse of current is then

$$H(t) = \int_{-\infty}^{\infty} \left(\frac{k}{i\omega + \gamma} + \frac{\bar{k}}{i\omega + \bar{\gamma}} \right) e^{i\omega t} df = ke^{-\gamma t} + \bar{k}e^{-\bar{\gamma}t}, \quad t > 0. \quad (27)$$

To find the envelope function of the voltage drop when the frequency modulated current

$$I(t) = \frac{h}{2} (e^{i\omega_0 t + i\theta(t)} + e^{-i\omega_0 t - i\theta(t)}) \quad (1)$$

is applied to the circuit, we make use of the general formula, (17), which states that the spectrum, or Fourier transform, of this envelope function is

$$F_0(f) = \int_{-\infty}^{\infty} [a(t) + ib(t)] e^{-i\omega t} dt = \Psi(f) Z(f + f_0) \quad (17)$$

where $\Psi(f)$ is the F.T. of $h \exp [i\theta(t)]$ and $Z(f)$ is the impedance, (26), in this case. The envelope function then is

$$\begin{aligned} a(t) + ib(t) &= \int_{-\infty}^{\infty} \Psi(f)Z(f + f_0)e^{i\omega t} df \\ &= hke^{-(i\omega_0 + \gamma)t} \int_{-\infty}^t e^{(i\omega_0 + \gamma)\tau + i\theta(\tau)} d\tau \\ &\quad + h\bar{k}e^{-(i\omega_0 + \bar{\gamma})t} \int_{-\infty}^t e^{(i\omega_0 + \bar{\gamma})\tau + i\bar{\theta}(\tau)} d\tau. \end{aligned} \quad (28)$$

This result is obtained by employing the convolution formula⁵, which states that if F_1 and F_2 are F.T.'s of G_1 and G_2 , respectively, then the F.T. of F_1F_2 is

$$\int_{-\infty}^{\infty} F_1(f)F_2(f)e^{i\omega t} df = \int_{-\infty}^{\infty} G_1(\tau)G_2(t - \tau) d\tau. \quad (29)$$

(The upper limit of the integrals in (28) is t instead of ∞ for the reason that $H(t)$ is zero for $t < 0$.) The result could have been obtained equally well without using the Fourier transforms by substituting (27) in (8). When $\theta(t)$ is specified mathematically in the infinite interval $(-\infty, \infty)$ these formulas give the resultant of the steady state and transient oscillations. Various problems can be solved by specifying particular forms of variation for $\theta(t)$. Two examples follow: (1) where the instantaneous frequency, $\theta'(t)$, is an impulse and (2), where $\theta'(t)$ is a sinusoidal wave, as for elementary signal transmission.

Example 1: Impulse Modulation

Ignition interference comprises a sequence of sharp impulses, each of duration very brief compared to the interval between them, so that the transient in the receiver produced by one impulse dies away before the next one arrives. It is therefore sufficient to consider the disturbance caused by a single impulse.

If the receiver is perfectly tuned, an impulse produces, in the tuned circuits, a transient of the same nominal frequency⁶ as the signal carrier. When superposed on the carrier, the interfering transient alters, or modulates, both the amplitude and phase of the carrier. The amplitude modulation is wiped out by the limiter, but the phase modulation remains to produce noise in the output. The phase shift caused by the transient is a random variable, because it depends upon the time of arrival of the impulse, and this is entirely fortuitous.

⁵ See pair 202 of Item 10 in the bibliography.

⁶ By "nominal frequency" is meant the frequency as determined by counting zeros of the wave. The transient actually comprises a spectrum of frequencies spread over the band of the tuned circuits, of course.

It is of engineering interest to determine the noise produced by a very large impulse, exceeding greatly the amplitude of the signal carrier. When such a large impulse arrives, it causes a sudden jump, or discontinuity, in the phase of the carrier. The excursion of the instantaneous frequency corresponding to the phase jump is indefinitely large and the problem accordingly cannot be solved satisfactorily by means of the usual assumption of quasi-stationary frequency. The problem of large impulsive interference provides, therefore, the principal justification for the more exact method of analysis here employed. In the paragraph following, the problem is restated in terms providing a suitable basis for mathematical analysis.

We assume, as before, that the limiter is delivering to the frequency detector a steady carrier current of constant amplitude h and frequency f_0 . At time $t = 0$ a brief disturbance occurs specified by the statement that the instantaneous frequency, $\theta'(t)$, of the current suddenly executes an impulse of moment Θ . That is: $\theta'(t)$ is zero at all times except at $t = 0$, when it goes to infinity and back to zero again in such a way that the area of the impulse so formed is Θ . The carrier current amplitude then remains constant but the phase, $\theta(t)$, of the carrier takes a sudden jump of Θ radians at $t = 0$. What is the voltage output of the frequency detector?

The general formula (28) gives directly the envelope function of the voltage across the impedance (26) for a phase function $\theta(t)$. In this formula we have now to put $\theta(t) = 0$ before time $t = 0$ and Θ , after $t = 0$. We do this by dividing the interval of integration into two parts, $(-\infty, 0)$ and $(0, t)$; thus

$$\begin{aligned} a(t) + ib(t) &= hke^{-(i\omega_0 + \gamma)t} \left(\int_{-\infty}^0 e^{(i\omega_0 + \gamma)\tau} d\tau + e^{i\Theta} \int_0^t e^{(i\omega_0 + \gamma)\tau} d\tau \right) \\ &\quad + h\bar{k}e^{-(i\omega_0 + \bar{\gamma})t} \left(\int_{-\infty}^0 e^{(i\omega_0 + \bar{\gamma})\tau} d\tau + e^{i\Theta} \int_0^t e^{(i\omega_0 + \bar{\gamma})\tau} d\tau \right) \\ &= hk \frac{(1 - e^{i\Theta})e^{-(i\omega_0 + \gamma)t} + e^{i\Theta}}{i\omega_0 + \gamma} + h\bar{k} \frac{(1 - e^{i\Theta})e^{-(i\omega_0 + \bar{\gamma})t} + e^{i\Theta}}{i\omega_0 + \bar{\gamma}}. \quad (30) \end{aligned}$$

Let the radian frequency interval by which the applied frequency ω_0 is set off from the resonant frequency ω_c be

$$\Delta = \omega_0 - \omega_c \quad (31)$$

as indicated on the curves of Fig. 2. When α/ω_c is small compared to unity, as it is in practical circuits, β is very nearly equal to ω_c . (See the formulas following equations (26).) Then

$$i\omega_0 + \gamma = i\omega_0 + \alpha + i\omega_c = \alpha - i\Delta + 2i\omega_0$$

and

$$i\omega_0 + \bar{\gamma} = i\omega_0 + \alpha - i\omega_c = \alpha + i\Delta. \quad (32)$$

From this it is evident that the first term of (30) contains the demodulation sum product of frequency on the order of $2\omega_0$. This frequency will be suppressed by the diode load circuit and consequently the second term of (30) is an adequate representation of the envelope function. Therefore we write

$$a(t) + ib(t) = h\bar{k} \frac{(1 - e^{i\theta})e^{-(i\omega_0 + \bar{\gamma})t} + e^{i\theta}}{i\omega_0 + \bar{\gamma}} \quad (33)$$

and with the above approximations this is very nearly equal to

$$a(t) + ib(t) = \frac{h}{2C} \frac{(1 - e^{i\theta})e^{-(\alpha + i\Delta)t} + e^{i\theta}}{\alpha + i\Delta}. \quad (34)$$

One deduction that can be made immediately from this formula is that the frequency of the oscillation in the output of the rectifier caused by the phase jump at the input is Δ , the radian frequency interval by which the applied carrier frequency differs from the resonant frequency. The oscillation is heavily damped, however, because α , while being very small compared to ω_c , is comparable in magnitude with Δ in circuits commonly used.

The angle of the complex envelope function (34) represents merely a phase shift of the carrier frequency ω_0 . We are interested only in the magnitude of the function, viz.:

$$c(t) = [a^2(t) + b^2(t)]^{1/2}. \quad (35)$$

After some algebraic work, the desired formula comes out of (34) in the following form:

$$c(t) = \frac{h}{2C} \frac{\left[1 - 2m(t) \sin\left(\Delta t + \frac{\theta}{2}\right) + m^2(t)\right]^{1/2}}{(\alpha^2 + \Delta^2)^{1/2}}, t > 0$$

where

$$m(t) = 2e^{-\alpha t} \sin \frac{\theta}{2} \quad (36)$$

The discussion so far has dealt with a single impedance (or network) and has been concerned with obtaining formulas for the voltage across the impedance, and the envelope thereof, when a frequency-modulated current is sent through it. It is necessary now to refer to the construction of the balanced frequency detector, which is the particular object of our study. Figure 1 shows two impedances having the variation with frequency sketched in Fig. 2. The carrier current is driven through the two impedances in series and linear rectifiers are connected across each in such polarity that their low-frequency output voltages are opposed. We assume that the output

voltage of each rectifier is the envelope of the voltage existing across its associated impedance. Therefore, to find the total output of the balanced frequency detector, we have to find the difference between the envelopes of these voltages.

It is necessary to specify the two impedances more precisely. It appears that the best operation is obtained if the frequency of the carrier is midway between the resonant frequencies of the two impedances. That is

$$\Delta = \omega_0 - \omega_1 = \omega_2 - \omega_0$$

where ω_1 , ω_2 are the resonant frequencies of Z_1 , Z_2 , (previously written as ω_c , for any impedance). Furthermore it appears that the two impedances should have identical values of C and very nearly the same damping constants. The design of the detector circuit is accordingly specified by

$$C_1 = C_2 = C$$

$$R_1 = R_2 = R$$

$$L_1 = 1/\omega_1^2 C$$

$$L_2 = 1/\omega_2^2 C$$

and then

$$\alpha_1 = \alpha_2 = \alpha = 1/2RC$$

$$k_1 = k_2 = k = 1/2C \quad (a/\omega_0 \ll 1) \quad (37)$$

$$\sqrt{\frac{L_1}{L_2}} = \frac{\omega_2}{\omega_1} = \frac{\omega_0 + \Delta}{\omega_0 - \Delta}.$$

All the quantities are assumed to be substantially constant over the significant frequency range.

With the circuit constants so proportioned, it can be seen from (36) that the envelope of the voltage across Z_1 differs from that across Z_2 only in the sign of Δ . Therefore, the output voltage of the balanced frequency detector, when the instantaneous frequency variation is an impulse of moment Θ at $t = 0$, is

$$\begin{aligned} V_0(t) &= c_2(t) - c_1(t) \\ &= \frac{h}{2C} (\alpha^2 + \Delta^2)^{-1/2} \left(\left[1 + 2m \sin \left(\Delta t - \frac{\Theta}{2} \right) + n^2 \right]^{1/2} \right. \\ &\quad \left. - \left[1 - 2m \sin \left(\Delta t + \frac{\Theta}{2} \right) + m^2 \right]^{1/2} \right), \quad t > 0 \\ &= 0, \quad t < 0, \quad \text{where } m = 2e^{-\alpha t} \sin \frac{\Theta}{2}. \end{aligned} \quad (38)$$

On Fig. 5 is given a plot of this function for a value $\alpha/\Delta = 1$ of the relative damping. Calculations for other values of α/Δ show that the output oscillates only weakly for $\alpha/\Delta = \frac{1}{2}$ and is nearly dead-beat for $\alpha/\Delta = 2$

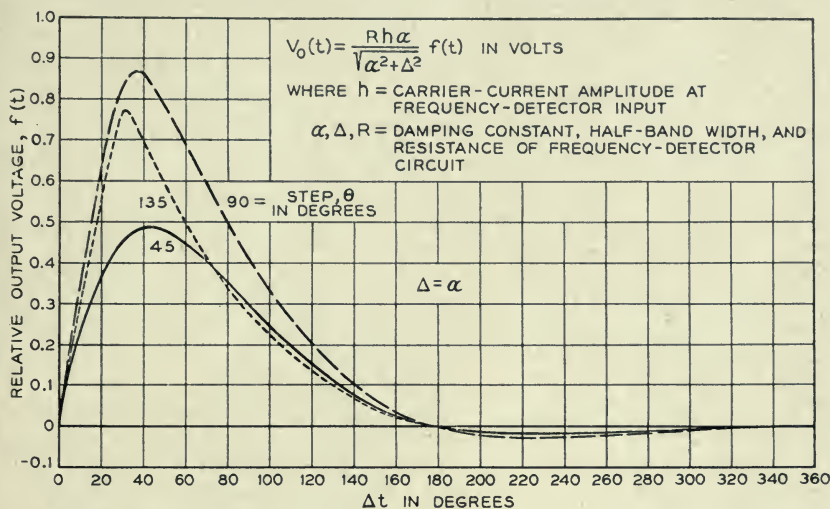


Fig. 5—Transient response of a balanced frequency detector for a step θ in the carrier phase.

Example 2: Signal Reception

If the carrier is frequency-modulated by a signal

$$s(t) = S \cos qt \quad (39)$$

and the frequency modulation factor of the transmitter is μ , the phase of the carrier wave is made to vary in accordance with the relation

$$\theta(t) = \mu \int_{-\infty}^t s(t) dt = \frac{\mu S}{q} \sin qt = x \sin qt \quad (40)$$

μS is then the radian frequency deviation of the transmitter and $\mu S/q$, which is the ratio of this frequency deviation to the frequency of the signal, is commonly referred to as the "frequency deviation ratio." This factor, which is denoted by x in the above equation, enters as a fundamental parameter in all FM theory.

To find the envelope function of the voltage wave produced across the impedance (26), when the frequency-modulated carrier is received at the frequency detector, we put (40) in (28). To effect the integration, the expansion

$$e^{ix \sin qt} = \sum_{n=-\infty}^{\infty} J_n(x) e^{inqt} \quad (41)$$

is used, $J_n(x)$ being the Bessel coefficient of the first kind, of the n th order and with the argument x . For brevity, $J_n(x)$ will be written J_n . Also, the first term of (28) is to be omitted, because, as was shown in the preceding example, it represents frequency sum terms which are filtered out by the diode output circuit. Then we have

$$a(t) + ib(t) = h\bar{k}e^{-(i\omega_0 + \bar{\gamma})t} \int_{-\infty}^t e^{(i\omega + \bar{\gamma})\tau} \sum_{n=-\infty}^{\infty} J_n e^{inq\tau} d\tau. \quad (42)$$

When the integration is carried out, the result is

$$a(t) + ib(t) = h\bar{k} \sum_{n=-\infty}^{\infty} \frac{J_n e^{inqt}}{i\omega_0 + \bar{\gamma} + inq} = h\bar{k} \sum_{n=-\infty}^{\infty} \frac{J_n e^{inqt}}{\alpha + i\Delta + inq}. \quad (43)$$

To obtain the magnitude of $a + ib$, which is the envelope required, we multiply the above Fourier series for $a + ib$ by that for $a - ib$, obtaining a double summation, which can be written as follows:

$$\begin{aligned} c^2(t) &= a^2(t) + b^2(t) = c_0 + \sum_{n=1}^{\infty} (c_n e^{inqt} + \bar{c}_n e^{-inqt}) \\ &= a_0 + 2 \sum_{n=1}^{\infty} a_n \cos nqt \end{aligned} \quad (44)$$

where a_n is the real part of the complex coefficient c_n and $a_0 = c_0$.

The coefficients are given by

$$\begin{aligned} c_n &= \frac{h^2}{4C^2} \sum_{m=-\infty}^{\infty} \frac{J_m J_{m+n}}{(\alpha - i\Delta - imq)[\alpha + i\Delta + i(m+n)q]} \\ \bar{c}_n &= \frac{h^2}{4C^2} \sum_{m=-\infty}^{\infty} \frac{J_m J_{m+n}}{(\alpha + i\Delta + imq)[\alpha - i\Delta - i(m+n)q]}. \end{aligned} \quad (45)$$

Obtaining this result involves use of the relation

$$J_{-n}(x) = (-)^n J_n(x). \quad (46)$$

From (45)

$$a_n = \frac{h^2}{4C^2} \sum_{m=-\infty}^{\infty} \frac{J_m J_{m+n} [\alpha^2 + (\Delta + mq)(\Delta + mq + nq)]}{[\alpha^2 + (\Delta + mq)^2][\alpha^2 + (\Delta + mq + nq)^2]}. \quad (47)$$

Finally, we have to obtain, as before, the difference between the envelopes of the voltages across the two impedances of the balanced frequency detector; that is, we have to determine

$$V_0(t) = c_2(t) - c_1(t) \quad (48)$$

where $c_1(t)$ is given by (44) as it stands and $c_2(t)$ is obtained from the same expression merely by reversing the sign of Δ . The complete solution for

the output voltage of the frequency detector, when a phase variation, $\theta'(t) = xq \cos qt$, is impressed on the carrier at the input, then is

$$V_0(t) = \left[a_0 + 2 \sum_{n=1}^{\infty} a_{n2} \cos nqt \right]^{1/2} - \left[a_0 + 2 \sum_{n=1}^{\infty} a_{n1} \cos nqt \right]^{1/2} \quad (49)$$

where a_{n1} is given by (47) and a_{n2} is given by the same formula with the sign of Δ reversed.

An approximation that is permissible when the frequency swing xq does not approach the available frequency range Δ is

$$V_0(t) = \sum_{n=1}^{\infty} A_n \cos nqt \quad (50)$$

where

$$A_n = \frac{a_{n2} - a_{n1}}{\sqrt{a_0}}, \quad n \text{ odd}; = 0, n \text{ even.} \quad (51)$$

The table following gives the results of a computation of the first four coefficients from formulas (45) and (51) for the case of a frequency deviation ratio, $x = 5$, for $q/2\pi = 3000$ cycles per second, for $\Delta/2\pi = 30,000$ cycles per second and for $\alpha/\Delta = \frac{1}{2}, 1$ and 2 .

Coefficient	$\alpha/\Delta = .5$	1.0	2.0
A_0	0	0	0
A_1	.848	.478	.191
A_2	0	0	0
A_3	.0312	-.00438	-.00281

Note: To obtain volts, multiply all values by $\frac{R h \alpha}{(\alpha^2 + \Delta^2)^{1/2}}$

The coefficients for even values of n vanish, which confirms what can be inferred from physical considerations, namely, that the balanced construction of the frequency detector eliminates the d.c. component and all even harmonics. From the ratio of A_3 to A_1 we obtain the following ratios, expressed in db's, of the third harmonic distortion to the fundamental signal for the three circuit designs:

α/Δ	$20 \log 10 A_3/A_1 $
.5	-28.7 db
1.0	-40.8
2.0	-36.6

The results for the sinusoidal signal, when considered in conjunction with those for the impulse modulation, also permit certain conclusions regarding signal-to-noise ratios for impulsive interference in FM reception.

Ratio of Noise to Signal

It may be helpful, in conclusion, to attempt a theoretical estimate of the ratio of noise to signal in the audio output under the condition of severe impulsive interference.

The ratio of the peak value of the pulse to the signal amplitude at the frequency detector output is given by the ratio of the peak values of $f(t)$, as plotted in figure 5, to the values of A_1 in the table above. To obtain a result of practical significance, however, the effect of the audio circuit should be taken into account. In the absence of specific information on the structure of this circuit, we assume that the peak value of a pulse at its output is equal to the area, or moment, of the pulse at its input times twice the audio cutoff frequency. This is true for an ideal "square cutoff" filter and not seriously in error for actual circuits. The area of the largest pulse at the frequency detector output is approximately $2\Delta/(\alpha^2 + \Delta^2)$ and the value of A_1 , the signal fundamental amplitude, can be approximated by

$$A_1 = \frac{2\Delta x q}{\alpha^2 + \Delta^2} \quad (52)$$

(For the example above, this approximation gives $A_1 = .8, .5$ and $.2$ as compared to the exact values, $.848, .478$ and $.191$.) In this way we arrive at the following estimate of the peak ratio of noise to signal in the audio output:

$$\frac{\text{Max. value of largest pulse}}{\text{Signal amplitude}} = \frac{\omega_a}{\pi x q} \quad (53)$$

where ω_a is the cutoff-frequency of the audio circuit, q the signal frequency and x the frequency deviation ratio. Then xq is the "frequency swing" of the transmitter, i. e., the maximum departure of the instantaneous frequency from its mean value. It is to be noted that this formula is free from the detector circuit parameters, α, Δ, R , and indicates that, to a first approximation, at least, the maximum ratio of noise to signal depends only upon the audio circuit cutoff frequency and the FM swing. Furthermore, this establishes a ceiling for the interference that will not be exceeded no matter how large the impulses may be.

BIBLIOGRAPHY

1. Variable Frequency Electric Circuit Theory with Application to the Theory of Frequency-Modulation, John R. Carson and Thornton C. Fry in *Bell System Technical Journal*, Vol. XVI, No. 4, October 1937.
2. The Detection of Frequency Modulated Waves, J. G. Chaffee in *Proc. I.R.E.*, Vol. 23, May 1935. (*Bell Tel. Sys. Monograph B-863*)
3. Effects of Tuned Circuits upon a Frequency Modulated Signal, Hans Roder in *Proc. I.R.E.*, Vol. 25, December 1937.
4. The Reception of Frequency-Modulated Radio Signals, Victor J. Andrew in *Proc. I.R.E.*, Vol. 20, May 1932.

5. The Phase Discriminator, K. R. Sturley in *Wireless Engineer*, February 1944.
6. Radio Engineers Handbook, F. E. Terman, First Ed., pp. 585-588.
7. Motor Car Ignition Interference, C. C. Eaglesfield in *Wireless Engineer*, 23, 1946.
8. Interference Problems in Frequency Modulation, F. L. H. M. Stumpers in *Philips Research Reports*, 2, 1947.
9. On the Calculation of Impulse Noise Transients in Frequency Modulation Receivers, F. L. H. M. Stumpers in *Philips Research Reports*, 2, 1947.
10. Fourier Integrals for Practical Applications, George A. Campbell and Ronald M. Foster in *Bell System Technical Journal*, October 1928—Monograph B-584.

Transverse Fields in Traveling-Wave Tubes

By J. R. PIERCE

Traveling-wave tubes will have gain even if the r-f field at the mean position of the electron stream is purely transverse. The addition of a longitudinal magnetic focusing field reduces the gain due to transverse fields and increases the electron velocity for optimum gain.

ALL slow electromagnetic waves have both longitudinal and transverse electric field components. Sometimes either the longitudinal or the transverse field may go to zero along a line or plane parallel to the direction of propagation. For instance, for the slow mode of propagation there is no transverse field on the axis of a helically-conducting sheet. Still, over any plane normal to the direction of propagation there are bound to be both longitudinal and transverse field components.

If a very strong longitudinal magnetic field is used in connection with a traveling-wave tube, the transverse motions of electrons may be so restricted as to be of little importance. With weak focusing fields, however, the transverse motion of electrons may be important in producing gain. The transverse fields can force the electrons sidewise, and thus change the longitudinal fields acting on them in such a way as to abstract energy from the electron stream.¹ This is closely analogous to the action of the longitudinal fields in displacing electrons forward or backward into regions of greater or lesser longitudinal field.

The purpose of this paper is to analyze the behavior of traveling-wave tubes in which transverse fields are important. The attack will be similar to that used previously.²

1. CIRCUIT THEORY

In this paper we shall consider only the electric field associated with the slow mode of propagation along the circuit having a speed close to the electron speed, and we shall neglect other field components attributable to local space charge. The writer believes the results so obtained to be valid at low currents but in error at high currents, and an acceptable guide at currents usually encountered.

In an earlier paper² a relation was found between the longitudinal field E_z excited in a mode of propagation of a transmission system and the longitud-

¹ See, for instance, J. R. Pierce and W. G. Shepherd, "Reflex Oscillators," *B. S. T. J.*, Vol. 26, No. 3, pp. 666-670 (July, 1947).

² J. R. Pierce, "Theory of the Beam-Type Traveling-Wave Tube," *Proc. I. R. E.*, Vol. 35, pp. 111-123, Feb. 1947.

inal exciting current q . Both E_z and q vary as $(\exp j\omega t) (\exp -\Gamma z)$. The relation is

$$E_z = q \frac{\Gamma_0}{\psi_0^* (\Gamma^2 - \Gamma_0^2)}. \quad (1)$$

Here Γ_0 is the propagation constant of the transmission mode considered and is defined in such a sense that for unattenuated propagation, $\Gamma_0 = j\beta_0$ where β_0 is a positive number. The quantity ψ_0 is defined as

$$\psi_0 = \frac{2P}{E_z E_z^*}. \quad (2)$$

Here P is complex power transmitted by the mode and E_z is the field associated with the mode.

In generalizing (1), let us consider the combination of equations (1) and (2)

$$P^* = \frac{1}{2} q E_z^* \frac{\Gamma_0}{\Gamma^2 - \Gamma_0^2}. \quad (3)$$

Now, suppose there is motion of the electrons not only in the z direction but in a direction normal to the z direction, which we will call the y direction. We shall have two extra first-order terms of the same general nature as $q E_z^*$, which contribute to the power, giving

$$P^* = \frac{1}{2} \left(q_z E_z^* + (-I_0) y \frac{\partial E_z^*}{\partial y} + q_y E_y^* \right) \frac{\Gamma_0}{\Gamma^2 - \Gamma_0^2}. \quad (4)$$

Here q_z is the a-c convection current in the z direction, $-I_0$ is the d-c convection current in the z direction (assumed to be the only d-c convection current), y is a small displacement, q_y is the convection current in the y direction and E_y is the field in the y direction.

We will now specialize this expression. Suppose we consider a two-dimensional transverse magnetic wave propagating in the z direction with a phase velocity v such that $v^2 \ll c^2$. Then over a restricted region the electric field can be represented quite accurately as the gradient of a scalar potential of

$$V = \exp(-\Gamma z) (A \exp(j\Gamma y) + B \exp(-j\Gamma y)). \quad (5)$$

Here A and B are constants. Using our notation, in which the field is understood to include the factor $\exp(-\Gamma z)$, we obtain

$$E_z = \Gamma (A \exp(j\Gamma y) + B \exp(-j\Gamma y)) \quad (6)$$

$$\frac{\partial E_z}{\partial y} = j\Gamma^2 (A \exp(j\Gamma y) - B \exp(-j\Gamma y)) \quad (7)$$

$$E_y = -j\Gamma (A \exp(j\Gamma y) - B \exp(-j\Gamma y)). \quad (8)$$

In other words

$$\frac{\partial E_z}{\partial y} = -\Gamma E_y. \quad (9)$$

This relation will also be approximately correct remote from the axis in an axially symmetrical tube. Here we let y represent a displacement in the r direction.

We may also define a quantity α so that

$$E_y = j\alpha E_z \quad (10)$$

$$\alpha = \frac{-(A \exp(j\Gamma y) - B \exp(-j\Gamma y))}{Ay \exp(j\Gamma y) + B \exp(-j\Gamma y)}. \quad (11)$$

For an active mode, such as the one we consider, the chief component of $j\Gamma$ is a positive real number. Hence, for large positive values of y , the quantity α approaches a value

$$\alpha = 1. \quad (12)$$

This is characteristic of a plane symmetrical field far from the axis and also of an axially symmetrical field far from the axis.

Using (9) and (12) we rewrite (4)

$$P^* = \frac{1}{2}E_z^*[q_z - j\alpha^*(\Gamma^* I_0 y + q_y)] \quad (13)$$

We see from this that, according to our assumptions, for the mode considered,

$$E_z = (q_z - j\alpha^*(\Gamma^* I_0 y + q_y)) \frac{\Gamma_0}{\psi_0^*(\Gamma^2 - \Gamma_0^2)}. \quad (14)$$

We will henceforward assume that α and ψ are so nearly real that we can regard them as real quantities, giving

$$E_z = [q_z - j\alpha(\Gamma^* I_0 y + q_y)] \frac{\Gamma_0}{\psi_0(\Gamma^2 - \Gamma_0^2)}. \quad (15)$$

This is, then, the circuit equation which we will use.

2. ELECTRONICS EQUATIONS

We will assume an unperturbed motion of velocity u_0 in the z direction, parallel to a uniform magnetic field of strength B . Products of a-c quantities will be neglected.

In the x direction, perpendicular to the y and z direction

$$\frac{d\dot{x}}{dt} = -\eta B \dot{y}. \quad (16)$$

Assume that $\dot{x} = 0$ at $y = 0$. Then

$$\dot{x} = -\eta B y. \quad (17)$$

In the y direction we have

$$\frac{d\dot{y}}{dt} = \eta(B\dot{x} - j\alpha E_z). \quad (18)$$

Now

$$\frac{d\dot{y}}{dt} = \frac{\partial \dot{y}}{\partial t} + \frac{\partial \dot{y}}{\partial z} \frac{dz}{dt} \quad (19)$$

$$\frac{d\dot{y}}{dt} = u_0(j\beta - \Gamma)\dot{y} \quad (20)$$

$$\beta = \frac{\omega}{u_0}. \quad (21)$$

We obtain from (20) and (18), and (17)

$$(j\beta - \Gamma)\dot{y} = -u_0\beta_0^2 y - \frac{j\eta\alpha E_z}{u_0} \quad (22)$$

$$\beta_0 = \frac{\eta B}{u_0}. \quad (23)$$

We may note that ηB is the electron cyclotron frequency. Now,

$$\dot{y} = \frac{\partial y}{\partial t} - \frac{\partial y}{\partial z} \frac{dz}{dt} \quad (24)$$

$$\dot{y} = u_0(j\beta - \Gamma)y.$$

From (24) and (22) we obtain

$$y = \frac{-j\eta\alpha E_z}{u_0^2[(j\beta - \Gamma)^2 + \beta_0^2]} \quad (25)$$

$$\dot{y} = \frac{-j\eta\alpha(j\beta - \Gamma)E_z}{u_0[(j\beta - \Gamma)^2 + \beta_0^2]}. \quad (26)$$

We will have for q_y

$$q_y = -I_0 \frac{\dot{y}}{u_0} \quad (27)$$

$$q_y = \frac{j\eta\alpha I_0(j\beta - \Gamma)E_z}{u_0^2[(j\beta - \Gamma)^2 + \beta_0^2]}. \quad (28)$$

It is easily shown that

$$\dot{z} = \frac{\eta E_z}{u_0(j\beta - \Gamma)}. \quad (29)$$

If ρ_0 is the d-c linear charge density and ρ the a-c linear charge density

$$\rho_0 = -\frac{I_0}{u_0}. \quad (30)$$

If q_z is the z component of convention current, we have

$$\begin{aligned} q_z &= \rho_0 \dot{z} + u_0 \rho \\ &= -\frac{I_0 \dot{z}}{u_0} + u_0 \rho. \end{aligned} \quad (31)$$

We also have

$$\begin{aligned} \frac{\partial q_z}{\partial z} &= -\frac{\partial \rho}{\partial t} \\ \Gamma q_z &= j\omega \rho. \end{aligned} \quad (32)$$

From (31) and (32) we obtain

$$q_z = \frac{-j\beta I_0 \dot{z}}{(j\beta - \Gamma)}. \quad (33)$$

Thus

$$q_z = \frac{j\eta\beta I_0 E_z}{u_0^2(j\beta - \Gamma)^2}. \quad (34)$$

3. COMBINED EQUATION

Combining (34), (28) and (25) with (15), we obtain

$$1 = \frac{\eta I_0}{u_0^2} \frac{\Gamma_0}{\psi_0(\Gamma^2 - \Gamma_0^2)} \left[\frac{j\beta}{(j\beta - \Gamma)_0^2} - \frac{\alpha^2(\Gamma^* - (j\beta - \Gamma))}{[(j\beta - \Gamma) + \beta_0^2]} \right]. \quad (35)$$

We now introduce new parameters

$$K = \frac{1}{\beta^2 \psi_0} = \frac{E_z E_z^*}{2\beta^2 P} \quad (36)$$

$$C^3 = \left(\frac{KI_0}{4V_0} \right). \quad (37)$$

Here P_0 is the power transmitted by the circuit for a field strength E_z . K has the dimensions of impedance. V_0 is the voltage specifying the electron velocity u_0

$$u_0^2 = 2\eta V_0. \quad (38)$$

From (36)–(38) and (35) we see

$$1 = \frac{2\beta^2 C^3 \Gamma_0}{(\Gamma^2 - \Gamma_0^2)} \left[\frac{j\beta}{(j\beta - \Gamma)^2} - \frac{\alpha^2(\Gamma^* - (j\beta - \Gamma))}{[(j\beta - \Gamma)^2 + \beta_0^2]} \right]. \quad (39)$$

We now make the approximation that

$$-\Gamma = -j\beta + \delta. \quad (40)$$

Where $|\delta| \ll |\beta|$. Neglecting higher order terms,

$$\frac{\Gamma^2 - \Gamma_0^2}{\Gamma_0} = 2j\beta^3 C^3 \left[\frac{1}{\delta^2} + \frac{\alpha^2}{\delta^2 + \beta_0^2} \right]. \quad (41)$$

4. PURELY TRANSVERSE FIELD ALONG THE PATH

We can imagine a case in which α approaches infinity and the quantity

$$D^3 = \alpha^2 C^3 = \frac{E_y E_y^*}{\beta^2 P} \frac{I_0}{8V_0} \quad (42)$$

remains finite. In this case we have

$$\frac{\Gamma^2 - \Gamma_0^2}{\Gamma_0} = \frac{2j\beta^3 D^3}{\delta^2 + \beta_0^2}. \quad (43)$$

We will let

$$-\Gamma_0 = -j\beta - j\beta D b - \beta D d. \quad (44)$$

Here b is a parameter describing the difference in speed between the electrons and the unperturbed wave and d is a loss parameter.

Assuming $bD \ll 1$ and $dD \ll 1$, and letting $\beta D(x + jy) = \delta$, we find

$$(x^2 - y^2 + f^2)(y + b) + 2xy(x + d) = -1 \quad (46)$$

$$(x^2 - y^2 + f^2)(x + d) - 2xy(y + b) = 0 \quad (47)$$

where

$$f^2 = \frac{\beta_0^2}{\beta^2 D^2}. \quad (48)$$

If would be difficult to work with all of the parameters b, f and d . However, it scarcely seems that the attenuation parameter d should enter into any unusual phenomena due to the presence of the magnetic field. Accordingly, let us investigate (46) and (47) for $d = 0$. We then obtain

$$x^2(3y + b) + (f^2 - y^2)(y + b) = -1 \quad (46a)$$

$$x[x^2 + (f^2 - y^2) - 2y(y + b)] = 0. \quad (47a)$$

From the $x = 0$ solution of (47a) we obtain

$$x = 0 \quad (49)$$

$$b = \frac{1}{y^2 - f^2} - y. \quad (50)$$

If it is found that this solution obtains for large and small values of b . For very large and very small values of b , either $y \doteq -b$ (51) or $y \doteq \pm f$ (52). The wave given by (50) is a *circuit* wave; that given by (51) represents the travel down the tube of electrons oscillating in the magnetic field with cyclotron frequency.

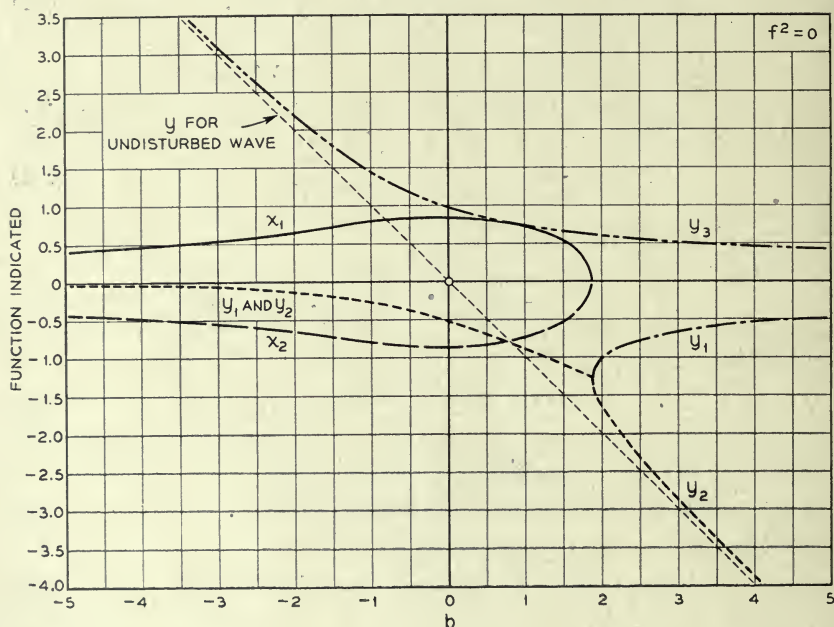


Fig. 1—Plot of parameters giving velocity and attenuation of the three forward waves vs. a parameter b proportional to electron speed with respect to the undisturbed wave. A positive value of x means an increasing wave; a positive value of y means a wave traveling faster than the electrons. This plot is for $f^2 = 0$ (no magnetic field).

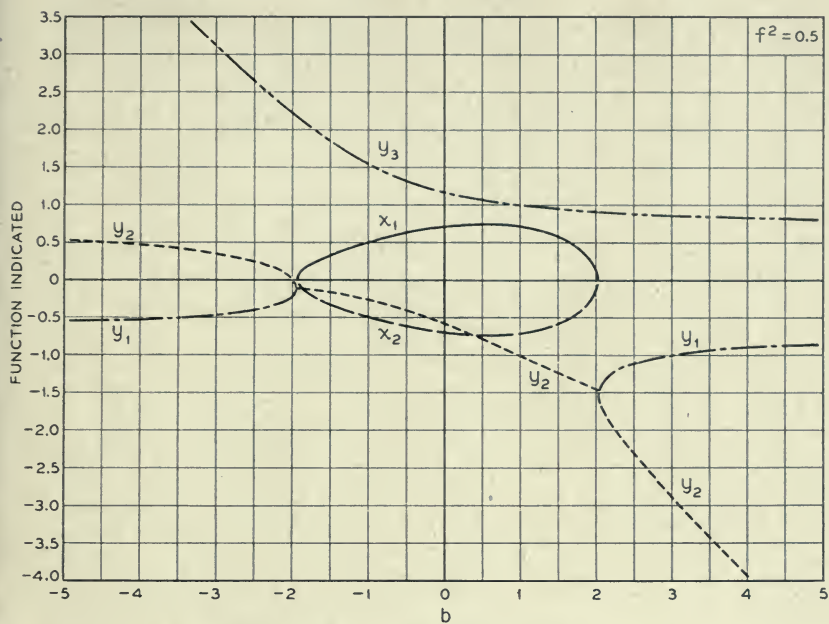
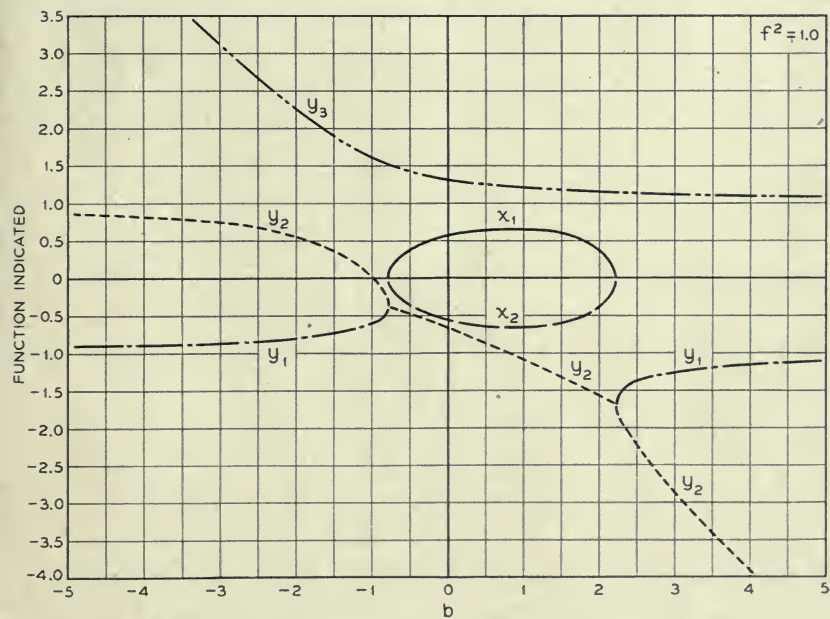
In an intermediate range of b , we have from (47a)

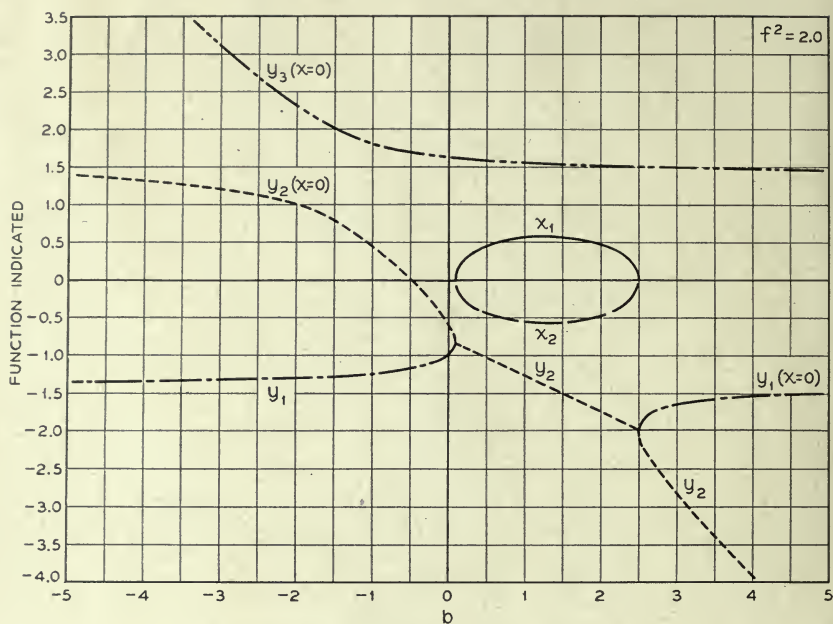
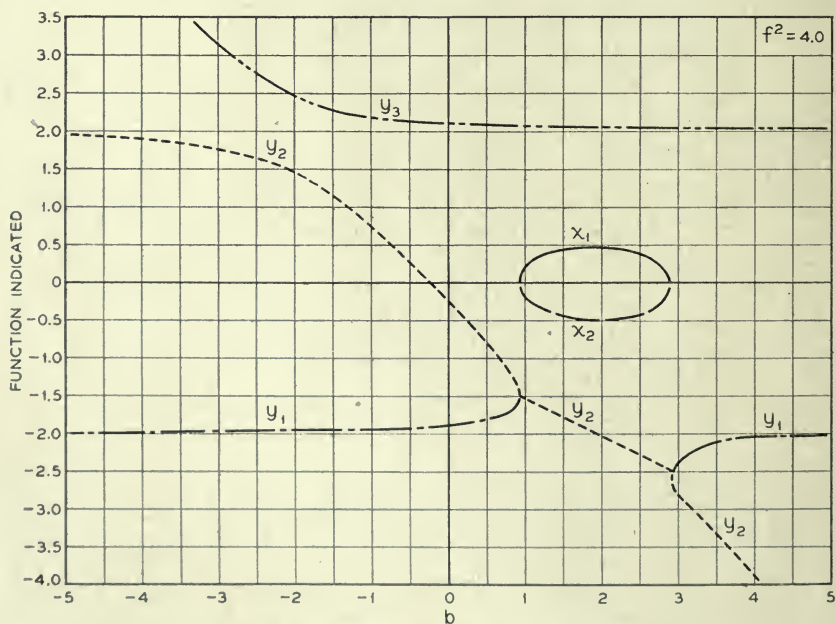
$$x = \pm \sqrt{2y(y + b) - (f^2 - y^2)} \quad (53)$$

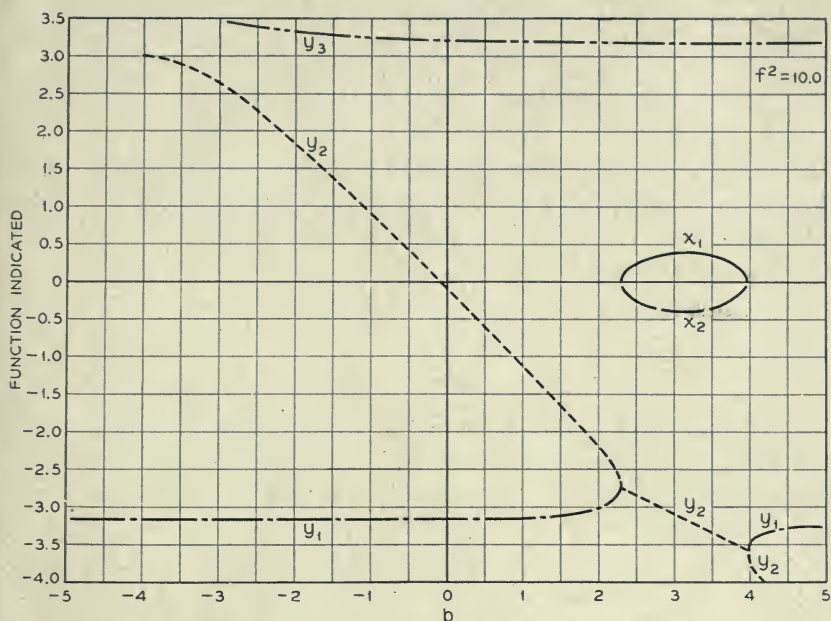
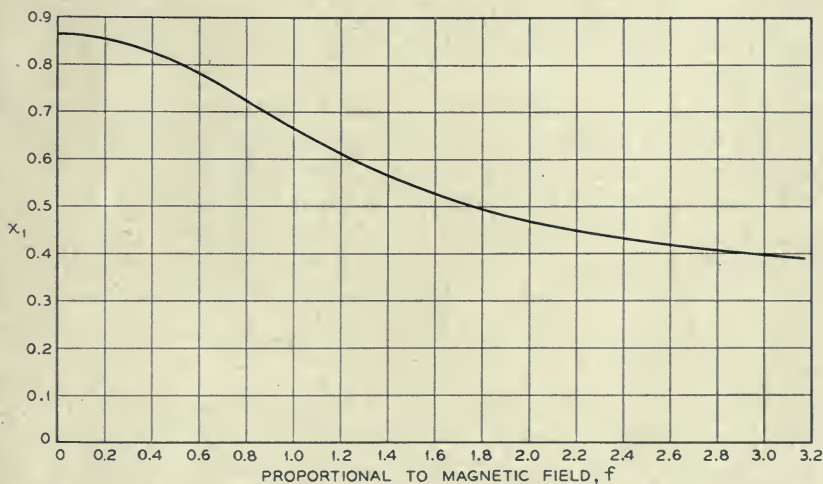
and

$$b = -2y \pm \sqrt{f^2 - 1/2y}. \quad (54)$$

For a given value of f^2 we can assume values of y and obtain values of b . Then, x can be obtained from (52) or (53). In Figs. 1–6, x and y are plotted vs. b for $f^2 = 0, .5, 1, 2, 4$ and 10 . It should be noted that x_1 , the parameter expressing the rate of increase of the increasing wave, has a maximum at larger values of b as f is increased (as the magnetic focusing field is increased). Thus, for higher magnetic focusing fields the electrons must be shot into the

Fig. 2—Propagation parameters for $f^2 = 0.5$.Fig. 3—Propagation parameters for $f^2 = 1.0$.

Fig. 4—Propagation parameters for $f^2 = 2.0$.Fig. 5—Propagation parameters for $f^2 = 4.0$.

Fig. 6—Propagation parameters for $f = 10.0$.Fig. 7—Parameter x giving rate of increase of increasing wave vs. f , which is proportional to magnetic field strength.

circuit faster to get optimum results than for low fields. In Fig. 7, the maximum positive value of x is plotted vs. f . The plot serves to illustrate the effect on gain of increasing the magnetic field.

Let us consider an example. Suppose

$$\lambda = 7.5 \text{ cm}$$

$$D = .03$$

These values are chosen because there is a longitudinal field tube which operates at 7.5 cm with a value of C (which corresponds to D) of about .03. The table below shows the ratio of the maximum value of x_1 to the maximum value of x_1 for no magnetic field.

Magnetic Field in Gauss	f	x_1/x_{10}
0	0	1
50	1.17	.71
100	2.34	.50

A field of 50 to 100 gauss should be sufficient to give useful focusing action. Thus, it may be desirable to use magnetic focusing fields in deflection traveling wave tubes. This will be more especially true in low-voltage tubes, for which D may be expected to be higher than .03.

5. MIXED FIELDS

In tubes designed for use with longitudinal fields, the transverse fields far off the axis approach in strength the longitudinal fields. The same is true of transverse field tubes far off the axis. Thus, it is of interest to consider equation (41) for cases in which α is neither very small nor very large, but rather is of the order of unity.

If the magnetic field is very intense so that β_0^2 is large, then the term containing α^2 , which represents the effect of transverse fields, will be very small and the tube will behave much as if the transverse fields were absent.

Consideration of both terms presents considerable difficulty as (41) leads to 5 waves (5 values of δ) instead of 3. The writer has attacked the problem only for the special case of $b = d = 0$. In this case we obtain from (41)

$$\delta = -j\beta^3 C^3 \left[\frac{1}{\delta^2} + \frac{\alpha^2}{\delta^2 + \beta_0^2} \right]. \quad (55)$$

In work which is given as an appendix, Dr. L. A. MacColl has shown that the two "new" waves (waves introduced when $\alpha = 0$) are unattenuated and thus unimportant and uninteresting (unless, as an off-chance, they have some drastic effect in fitting the boundary conditions).

Proceeding from this information, we will find the change in δ as β_0^2 is increased from zero. From (51) we obtain

$$d\delta = -j\beta^3 C^3 \left[-\frac{2d\delta}{\delta^3} - \frac{2\alpha^2 \delta d\delta}{(\delta^2 + \beta_0^2)^2} - \frac{2\alpha^2 d\beta_0^2}{(\delta^2 + \beta_0^2)^2} \right]. \quad (56)$$

Now, if $\beta_0 = 0$

$$-j\beta^3 C^3 = \frac{\delta^3}{(1 + \alpha^2)}. \quad (57)$$

Using this in connection with (56) for $\beta_0 = 0$ we obtain

$$d\delta = -\frac{2\alpha^2}{3\delta} d\beta_0^2. \quad (58)$$

For the increasing wave

$$\delta_1 = \beta(.866 - j.5). \quad (59)$$

Hence, for this wave

$$d\delta_1 = \frac{2}{3}(-.866 + j.5) \frac{d\beta_0^2}{\beta}. \quad (60)$$

This shows that applying a small magnetic field tends to decrease the gain. This does not mean, however, that the gain with a longitudinal and a transverse field and a magnetic field is less than the gain with the longitudinal field alone. To see this, we can assume that not β_0 but α^2 is small. Differentiating (55) we obtain

$$d\delta = -j\beta^3 C^3 \left[\frac{-2d\delta}{\delta^3} - \frac{2\alpha^2 \delta d\delta}{(\delta^2 + \beta_0^2)^2} - \frac{d\alpha^2}{(\delta^2 + \beta_0^2)} \right]. \quad (61)$$

If $\alpha = 0$

$$-j\beta^3 C^3 = \delta^3 \quad (62)$$

$$d\delta = \frac{1}{3} \frac{\delta^3 d\alpha^2}{(\delta^2 + \beta_0^2)}. \quad (63)$$

If β_0^2 is zero, a small transverse field (small increase in α^2) increases the magnitude of δ without changing the phase angle. If $\beta_0^2 \gg |\delta|^2$, then

$$d\delta = \frac{-j\beta^3 C^3}{\beta_0^2} d\alpha^2 \quad (64)$$

and the change in δ is purely imaginary. For the increasing wave, the change in δ as a transverse field is added will range from an increase in the real part for small magnetic fields to no change in the real part for large magnetic fields.

APPENDIX

STUDY OF THE ALGEBRAIC EQUATION

$$\delta = -j\beta^3 C^3 \left[\frac{1}{\delta^2} + \frac{\alpha^2}{\delta^2 + \beta_0^2} \right] \quad (1A)$$

$$\delta^3(\delta^2 + \beta_0^2) + j\beta^3 C^3(\delta^2 + \beta_0^2 + \alpha^2 \delta^2) = 0$$

$$\delta^5 + \beta_0^2 \delta^3 + j\beta^3 C^3(1 + \alpha^2)\delta^2 + j\beta_0^2 \beta^3 C^3 = 0$$

$$\left(\frac{\delta}{\beta_0}\right)^5 + \left(\frac{\delta}{\beta_0}\right)^3 + j\left(\frac{\beta}{\beta_0}\right)^3 C^3(1 + \alpha^2) \left(\frac{\delta}{\beta_0}\right)^2 + j\left(\frac{\beta}{\beta_0}\right)^3 C^3 = 0 \quad (2A)$$

Write

$$\begin{aligned} \frac{\delta}{\beta_0} &= z \\ \left(\frac{\beta}{\beta_0}\right)^3 C^3 &= a \\ \left(\frac{\beta}{\beta_0}\right)^3 C^3 a^2 &= b \end{aligned} \quad (3A)$$

Then

$$z^5 + z^3 + j(a + b)z^2 + ja = 0 \quad (4A)$$

a is assumed to be positive, and b is assumed to be real and non-negative. For $b = 0$ we have

$$(z^3 + ja)(z^2 + 1) = 0 \quad (5A)$$

$$z = j, -j, ja^{1/3}, ja^{1/3} e^{2\pi j/3}, ja^{1/3} e^{4\pi j/3} \quad (6A)$$

We have

$$\begin{aligned} [5z^4 + 3z^2 + 2j(a + b)z] \frac{\partial z}{\partial b} + jz^2 &= 0 \\ \frac{\partial z}{\partial b} &= \frac{-jz^2}{5z^4 + 3z^2 + 2j(a + b)z} \end{aligned} \quad (7A)$$

From this we draw the following conclusion. Suppose that for a certain value of b the five roots are distinct, and that among them there is a purely imaginary root. Then as b varies, in the neighborhood of its initial value, that root remains purely imaginary.

In particular, consider b as increasing from the initial value 0. As long as the five roots remain distinct, there are exactly three purely imaginary roots.

In order to have a real root $z = x$, we would have to have simultaneously

$$\begin{aligned} x^5 + x^3 &= 0 \\ (a + b)x^2 + a &= 0 \end{aligned} \quad (8A)$$

This is impossible (with $a > 0$). Hence there is never a real root.

In particular, as b increases from 0, no root can cross the real axis. Hence, as b increases from 0, as long as the roots remain distinct, there are two purely imaginary roots above the real axis, one purely imaginary root below the real axis, and two complex roots below the real axis.

Since there is no term in z^4 in the equation, the sum of all the roots is 0. Hence the two complex roots must be located symmetrically with respect to the imaginary axis.

First order variations of the roots with b can be calculated at once by means of the equation

$$\frac{\partial z}{\partial b} = \frac{-jz^2}{5z^4 + 3z^2 + 2j(a+b)z} \quad (9A)$$

In principle, higher-order variations can be calculated by carrying the differentiation to higher orders. However, the formulae get wonderfully complicated.

A very practical way of solving the equation is the following:

The three imaginary roots can be found by plotting a curve. If we let $z = jy$, (4A) becomes

$$y^5 - y^3 + (a+b)y^2 - a = 0 \quad (10A)$$

For the imaginary roots y is real and we have merely to plot the left-hand side of (10A) vs. y to find the roots. Denote them by z_1, z_2, z_3 , which are now regarded as known numbers. These roots satisfy the equation

$$(z - z_1)(z - z_2)(z - z_3) \equiv z^3 - (z_1 + z_2 + z_3)z^2 + (z_1z_2 + z_1z_3 + z_2z_3)z - z_1z_2z_3 \equiv z^3 + \alpha_1z^2 + \alpha_2z + \alpha_3 = 0 \quad (11A)$$

The two complex roots satisfy some equation

$$z^2 + \beta_1z + \beta_2 = 0 \quad (12A)$$

The β 's are at present unknown. When we find them we can at once calculate the complex roots. We must have

$$(z^3 + \alpha_1z^2 + \alpha_2z + \alpha_3)(z^2 + \beta_1z + \beta_2) \equiv z^5 + z^3 + j(a+b)z^2 + ja \quad (13A)$$

Comparing the coefficients of z^4 and z^0 , we get the equations

$$\begin{aligned} \alpha_1 + \beta_1 &= 0 \\ \alpha_3\beta_2 &= ja \end{aligned} \quad (14A)$$

which give us the β 's.

Suppose that the magnetic field is very small, so that $\beta_0 \ll \beta$. Then unless α is very small, both a and b in (10A) will be very large numbers, and we find that two of the imaginary roots are given approximately by

$$y = \pm \left(\frac{a}{a+b} \right)^{1/2} \quad (15A)$$

As $\beta_0 \rightarrow 0$ the other three roots are given by

$$z = [-j(a + b)]^{1/3} \quad (16A)$$

These three roots correspond to the waves found for traveling-wave tubes with a purely longitudinal field. The roots according to (15A) represent such a combination of deflection and bunching as to produce no induced current in the circuit. The roots of (15A) are "extra" roots attributable to the consideration of transverse fields and transverse electron motion.

For the roots given by (15A), $\delta/\beta \rightarrow 0$ as $\beta_0 \rightarrow 0$. Thus in this case it is convenient to form the solution of two parts, one varying as

$$e^{-j\beta z} \sin\left(\frac{a\beta_0}{a+b}\right)z$$

and the other varying as

$$e^{-j\beta z} \cos\left(\frac{a\beta_0}{a+b}\right)z$$

As $\beta_0 \rightarrow 0$, the first of these approaches the form

$$ze^{-j\beta z}$$

and the second approaches the form

$$e^{-j\beta z}$$

Again, these "extra" waves produce no induced current in the circuit.

Two additional pieces of information:

As $a \rightarrow 0$, b a remaining fixed, the roots approach the limiting values

$$0, 0, 0, j, -j.$$

As $a \rightarrow \infty$, ba remaining fixed, two of the roots approach the limiting values

$$\pm j \sqrt{\frac{a}{a+b}},$$

the other roots behave as

$$j(a+b)^{1/3}, j(a+b)^{1/3} e^{2\pi j/3}, j(a+b)^{1/3} e^{4\pi j/3}$$

Much of the preceding discussion depends upon the roots remaining distinct. The condition that two or more of the roots coincide, which is a relation between a and b , can be written out, but it has not as yet been reduced to a compact and intelligible form.

Abstracts of Technical Articles by Bell System Authors

*Pulse Code Modulation.*¹ H. S. BLACK and J. O. EDSON. A radically new modulation technique for multichannel telephony has been developed which involves the conversion of speech waves into coded pulses. An 8-channel system embodying these principles was produced. The method appears to have exceptional possibilities from the standpoint of freedom from interference, but its full significance in connection with future radio and wire transmission may take some time to reveal.

*Modulation in Communication.*² F. A. COWAN. Any signaling system requires some means for introducing a change in conditions at the sending end which may be recognized at the receiving end. The process by which the conditions are changed has come to be called modulation. There are many varieties of forms of change as well as a large number of conditions which are subject to change in response to the signals to be transmitted.

In early systems for communication at a distance the signal information might have been impressed upon a rising column of smoke, a light, a stone tablet, or a sheet of parchment. For modern communication wide use is made of systems in which the signals change the magnitude or condition of electric energy.

Starting with the electric telegraph a little more than a hundred years ago this medium of communication has grown steadily more important and more complex. To meet a variety of needs many systems of modulation have been developed and papers describing them in detail are available in the technical literature. Various aspects of the modulation processes have been analyzed carefully and presented and some of the earlier conceptions have acquired a classical textbook status. Recent trends have placed emphasis on modulation systems which more readily may be understood when viewed in a somewhat different manner. It is the purpose of this paper to present certain conceptions which may facilitate the understanding of the various systems of modulation and permit an improved perspective.

*Frequency Division Techniques for a Coaxial Cable Network.*³ R. E. CRANE, J. T. DIXON and G. H. HUBER. A description is given of developments employing frequency division techniques by which the telephone-message-carrying potentialities of the coaxial cable system are realized. By these methods 480 high quality telephone messages are prepared for

¹ *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 895-899).

² *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 792-796).

³ *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 1451-1459).

transmission over the line and restored to original condition at main terminal points. At intermediate points appropriate groups of channels may be removed, inserted, bridged, or relocated in the frequency spectrum of the line.

*Experimental Studies of a Remodulating Repeater.*⁴ W. M. GOODALL. This paper describes tests made on an experimental broad-band microwave f.m. repeater. A superheterodyne receiving unit is used with a microwave reflex-oscillator transmitting unit to form a repeater. An experimental setup for testing this repeater in a circulating-pulse system is described. Oscillograms showing the performance of the repeater on a multilink basis are discussed.

*An Electronic Regenerative Repeater for Teletypewriter Signals.*⁵ R. B. HEARN. The important factor in teletypewriter signal transmission over circuits is the relative position on a time scale of the code pulses. If this timing is preserved, wide amplitude variations can be experienced without errors resulting. Correctly timed signal pulses at the transmitting end of a circuit are not necessarily properly timed at the receiving end, as the transmission path may shift the timing of some transitions with respect to others. However, if the signals are not too badly changed or distorted, it is possible to retime them at an intermediate point and send them on in their original form.

Many arrangements have been devised for automatically retiming and retransmitting teletypewriter signals. These arrangements are known as regenerative repeaters. A few of these have been designed to make use of electronic timing arrangements and the purpose of this paper is to describe such an electronic regenerative repeater, known as repeater TG-29, designed originally for use by the Armed Forces.

*Submarine Detection by Sonar.*⁶ A. C. KELLER. Sonar, the only effective method of detecting completely submerged submarines was a major factor in the defeat of the U-boat and the winning of the Battle of the Atlantic. A majority of the 996 enemy submarines sunk during World War II was detected and located by sonar. The word sonar is formed from the phrase SOUNd Navigation And Ranging and applies broadly to under water sound devices for listening, echo ranging, and locating obstacles.

The QJA sonar system, one of those which got into active service during World War II, is described here. This equipment was designed by Bell Telephone Laboratories and manufactured by the Western Electric Company.

⁴ *Proc. I. R. E.*, May 1948 (pp. 580-583).

⁵ *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 904-911).

⁶ *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 1217-1230).

*Measurement of High Q Cavities at 10,000 Megacycles.*⁷ R. W. LANGE. Known methods of measuring Q in high Q resonant cavities, together with their accuracies and sources of error, are discussed. For relatively low values of Q and of frequency, it is shown that band width methods are more accurate than decrement methods. For values of Q above 30,000 at frequencies above 3,000 megacycles the reverse is true. The significant feature of the present method, the wide range heterodyne decrement method, is that the accuracy is improved by observing the decay over a relatively long interval of time. An absolute accuracy of plus or minus three per cent and a relative accuracy of plus or minus two per cent are achieved. Design features and performance are discussed and constructional details are presented.

*Absorbing Media for Underwater Sound Measuring Tanks and Baffles.*⁸ W. P. MASON and F. H. HIBBARD. By using absorbing walls surrounding a small body of water, measuring tanks have been produced which will determine the directional properties of underwater sound instruments down to a level of 25 db below the direct beam. These absorbing media are constructed by inserting fine mesh screen or packed copper wadding in a viscous liquid such as castor oil. These obstructions result in an enhanced viscous action which is nearly independent of the frequency above 10 kilocycles. A six-inch wall can reduce the reflections by 20 db. Tanks using such absorbing media were used for testing transducers at the manufacturing plant and were used for determining the approximate characteristics of small sized instruments. Absorbing media were also used in the sound transparent dome housing the transducer and in the back of the QJB transducers.

*Calculation of the Directivity Index for Various Types of Radiators.*⁹ C. T. MOLLOY. This paper gives the derivations of the "directivity index" formulas for several types of sound radiators. The "directivity index" is defined as "the ratio of the total acoustic power output of the radiator to the acoustic power output of a point source producing the same pressure at the same point on the axis." The utility of the directivity index concept is that it permits power calculations to be made for all radiators in the same manner as for point sources. Directivity index formulas, together with graphs covering practical cases, are given for the following types of radiators:

1. General plane piston in infinite baffle,
2. Circular plane piston in infinite baffle,
3. Rectangular plane piston in infinite baffle,
4. Sectoral horn,

⁷ *Trans. A. I. E. E.*, vol. 66, 1947 (pp. 161-166).

⁸ *Jour. Acous. Soc. America*, July 1948 (pp. 476-483).

⁹ *Jour. Acous. Soc. America*, July 1948 (pp. 387-405).

5. Multicellular horn,
6. Piston set in sphere.

*A Magnetic Field Strength Meter Employing the Hall Effect in Germanium.*¹⁰ G. L. PEARSON. The instrument to be described measures magnetic field strengths as determined from the Hall effect in germanium. The essential parts of this instrument include a small germanium probe, and a panel type microammeter calibrated directly in gauss. Its accuracy is ± 2 percent at fields between 100 and 8000 gauss. At higher fields the readings are too low, the error amounting to 9 percent at 20,000 gauss. The chief advantages of this instrument are: (a) small size and portability, (b) continuous reading rather than ballistic as in ordinary field strength meters, and (c) a small nonmagnetic probe with which one can search in very narrow gaps.

*The Representation of Vowels and their Movements.*¹¹ RALPH K. POTTER and GORDON E. PETERSON. It is shown that movement of the major resonances in the voiced sounds of speech may be represented by traces in a three-dimensional graph. Apparently a great deal can be learned about speech through investigation of such traces, and they suggest a new method for vowel designation that is particularly adaptable to quantitative analysis.

*General Mobile Telephone System.*¹² H. I. ROMNES and R. R. O'CONNOR. The tremendous need for communication with ships, airplanes, trucks, tanks, and other mobile units used in such large quantities during the war accelerated the development of practical mobile radiotelephone equipment for use in the 30 to 200 megacycle range and emphasized the practicability and usefulness of mobile telephone service. By the end of the war the art had advanced sufficiently in the applications of these higher frequencies so that it seemed practicable to provide telephone service to mobile units on a general basis rather than limit it to safety and emergency services as had been the case before the war. The Federal Communications Commission therefore made available a few frequencies for experiments in this field. In the two years which have elapsed, the Bell System has made this service available on an experimental basis in more than 60 cities and about 100 more systems are under construction. This paper describes the arrangements used and outlines the experience obtained to date with this service. Improvements are being made constantly so that this must be regarded as an interim report on a rapidly changing and expanding service.

*Interference between Very-High-Frequency Radio Communication Circuits.*¹³ W. RAE YOUNG, JR. Interference between different radio circuits is an old problem, one which in the past generally has been solved by trial and error

¹⁰ *Rev. Sci. Instruments*, April 1948 (pp. 263-265).

¹¹ *Jour. Acous. Soc. America*, July 1948 (pp. 528-535).

¹² *Trans. A. I. E. E.*, vol. 66, 1947 (1658-1666).

¹³ *Proc. I. R. E.*, *Waves and Electrons Section*, July 1948 (pp. 923-930).

and by hand tailoring (special filters, etc.). With the general increase in the usage of radio communication, however, the amount of potential interference is greatly increased. This paper will be concerned principally with the v.h.f. problem.

There is generally a large difference between transmitting and receiving power levels. As a result, spurious radiations, spurious responses, and lack of sufficient receiver selectivity may in many instances cause interference. Situations are described in which such interferences are likely.

In mobile systems interference can occur if the interfering station is close enough to "capture" it from the desired signal. This, in turn, depends upon the selectivity and spurious response of the receiver and the amount of spurious radiation from the transmitter. The problem can be approached in a statistical manner.

The types of spurious radio behavior which are common causes of interference are discussed. Sample measurements are given to illustrate the relative magnitude of the various modes of behavior. Formulas are given which permit computation of the frequency of the disturbances. A method is described for making charts suitable for a given type of equipment from which the spurious frequencies can be read directly as a function of the operating frequency.

Contributors to this Issue

J. T. MULLER, Technical College, Amsterdam, Holland, 1923; M.S., Stevens Institute of Technology, 1943. Mr. Muller came originally to the United States to continue his studies in 1923, became interested in the design and development of high-speed automatic equipment for a number of companies and joined Western Electric in 1941 and the Bell Telephone Laboratories in 1943. He has been primarily concerned with the protection of radar equipment against shock and vibration and the development of new types of shock mounts. His interest is in the field of experimental and mathematical dynamics.

WILLIAM W. MUMFORD, B.A., Willamette University, 1930. Bell Telephone Laboratories, 1930-. Mr. Mumford has been engaged in work that is chiefly concerned with ultra-short-wave and microwave radio communication.

LISS C. PETERSON, Chalmers Technical University, Gothenburg, 1921; Technical Universities of Charlottenburg and Dresden, 1921-23. American Telephone and Telegraph Company, 1925-30; Bell Telephone Laboratories, 1930-. Mr. Peterson has recently been concerned with the theory of hearing.

J. R. PIERCE, B.S. in Electrical Engineering, California Institute of Technology, 1933; Ph.D., 1936. Bell Telephone Laboratories, 1936-. Engaged in study of vacuum tubes.

CLAUDE E. SHANNON, B.S. in Electrical Engineering, University of Michigan, 1936; S.M. in Electrical Engineering and Ph.D. in Mathematics, M. I. T., 1940. National Research Fellow, 1940. Bell Telephone Laboratories, 1941-. Dr. Shannon has been engaged in mathematical research principally in the use of Boolean Algebra in switching, the theory of communication, and cryptography.

M. K. ZINN, B.S. in Electrical Engineering, Purdue University, 1918; U. S. Army Signal Corps and Air Service, 1918-19. American Telephone and Telegraph Company, Department of Development and Research, 1919-34; Bell Telephone Laboratories, 1934-. Mr. Zinn's work has been concerned mainly with transmission engineering problems.



